



UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTA' DI INGEGNERIA

Corso di Laurea Magistrale in Ingegneria Meccanica

Progettazione preliminare di macchine per il packaging: studio di metodi e strumenti di machine learning per la preventivazione dei costi

Preliminary design of packaging machineries: a study of machine learning methods and tools for cost estimation

Relatore:
Prof. Ing. Michele Germani

Tesi di laurea di:
Ing. Kevin Pignotti

Correlatore:
Ing. Federico Campi
Ing. Marco Mandolini

Anno Accademico 2019-2020

I. RINGRAZIAMENTI

Un ringraziamento al Professor Michele Germani, all'ingegner Domenico Polidori e ad ACMA per avermi dato l'opportunità di svolgere questo lavoro di tesi, ai Correlatori Federico Campi e Marco Mandolini nonché a Gianluca Porcelli, Tutor aziendale assegnatomi da ACMA, per avermi seguito e sostenuto lungo tutta la durata del tirocinio.

Un ringraziamento speciale a tutta la mia Famiglia senza la quale non sarei mai riuscito ad arrivare a questo punto.

Un ringraziamento a Priscilla che mi è stata sempre vicino, e ai miei amici Pierpaolo e Gian Marco con cui ho passato anni indimenticabili.

Infine auguro a me stesso di seguire sempre una logica positiva che mi porti ad un miglioramento continuo.

II. SOMMARIO

I.	Ringraziamenti.....	i
II.	Sommario.....	ii
III.	Elenco delle figure	v
IV.	Elenco delle tabelle	ix
0.	ACMA	1
0.1.	Generalità.....	1
0.2.	Storia	1
0.3.	Dati economici.....	3
1.	Introduzione.....	4
2.	Background scientifico.....	7
2.1.	Design to cost: progettare a costo target.....	7
2.2.	Tecniche per la stima dei costi	9
2.2.1.	Tecniche intuitive.....	11
2.2.2.	Tecniche analogiche	12
2.2.3.	Tecniche Analitiche	13
2.2.4.	Tecniche Parametriche	13
2.3.	Tecniche di machine learning.....	14
2.3.1.	Regressione lineare.....	15
2.3.2.	Regressione lineare semplice	16
2.3.3.	Regressione lineare multipla.....	17
2.3.4.	Artificial Neural Network	18
2.3.4.1.	Algoritmo di Backpropagation	20
2.3.4.2.	Cross validation.....	21
2.4.	Deep Learning	22
2.4.1.	Random Forest	23
2.4.2.	Gradient Boosting.....	25
2.5.	Studio di fabbricazione	27
2.5.1.	Metodi per effettuare uno studio di fabbricazione	28

2.5.2.	Pianificazione di processo	29
2.5.3.	Attività per la generazione del piano di processo	31
2.5.4.	Sistemi CAPP.....	32
3.	Stato dell'arte.....	35
3.1.	Strumenti per preventivazione dei costi.....	35
3.2.	Software sul mercato	36
3.2.1.	Software di costo CAD-based	36
3.2.2.	Software di costo generali	37
3.2.2.1.	SEER for Manufacturing	37
3.2.2.2.	MicroEstimating.....	38
3.2.2.3.	DFM Concurrent Costing e DFA Product Simplification	38
3.2.2.4.	Costimator.....	39
3.2.2.5.	aPriori Product Cost Management	39
3.2.3.	Strumenti di costo basati sul web.....	40
3.3.	Studi su modelli di costo parametrici	40
4.	Metodo	48
4.1.	Individuazione dei parametri a maggior sensibilità di errore	48
4.1.1.	Analisi dei costi "Should Cost"	51
4.1.2.	Software Utilizzato	54
4.1.2.1.	LeanCOST®	54
4.1.2.2.	LeanDESIGNER®	56
4.2.	Analisi modello parametrico.....	58
4.2.1.	Definizione Database per analisi parametrica.....	59
4.2.2.	Software Utilizzato	62
4.2.3.	Analisi dei risultati	64
5.	Caso Studio: Parte 1.....	65
5.1.	Preparazione del database	65
5.2.	Analisi efficienza globale	67
5.3.	Indagine sui parametri.....	70
5.3.1.	Analisi Pearson	70

5.3.2.	Studio del comportamento della variabile in funzione degli indici rimanenti.....	77
5.3.2.1.	LD base	77
5.3.2.2.	LD	78
5.3.2.3.	LeanCOST 11	86
5.3.2.4.	LeanCOST 9	92
5.3.2.5.	LeanCOST ACMA	98
5.4.	Tempistiche necessarie per la stima dei costi.....	102
6.	Caso Studio: Parte 2.....	104
6.1.	Preparazione Del Database.....	104
6.2.	Analisi Con Tool Sperimentale	105
6.3.	Risultati Del Modello Ottimale	108
6.3.1.	Tornitura	108
6.3.1.1.	Tornitura con parametri noti dal progettista	111
6.3.2.	Fresatura	113
6.3.2.1.	Fresatura con parametri noti dal progettista	116
6.3.3.	Convalida del modello.....	118
7.	Conclusioni.....	121
	Appendice A.....	126
8.	Bibliografia.....	127

III. ELENCO DELLE FIGURE

Figura 0.1 - Macchina automatica ACMA.....	2
Figura 0.2 - Logo ACMA S.p.A.....	3
Figura 2.1 - Andamento dei costi sostenuti e impegnati durante il ciclo di vita del prodotto.....	8
Figura 2.2 - Classificazione delle tecniche di stima dei costi.....	10
Figura 2.3 - Regressione lineare semplice	16
Figura 2.4 - Andamento coefficiente R^2	17
Figura 2.5 - Architettura reti neurali.....	19
Figura 2.6 - Cross validation	22
Figura 2.7 - Struttura Deep Learning.....	22
Figura 2.8 - Albero delle decisioni.....	24
Figura 2.9 - Struttura Random Forest	25
Figura 2.10 - Schema apprendimento Gradient Boosting.....	26
Figura 2.11 - L'operazione	30
Figura 2.12 - La sottofase	30
Figura 2.13 - Esempio di un piano di processo	31
Figura 2.14 - Albero delle decisioni.....	33
Figura 2.15 - Tabella delle decisioni	33
Figura 2.16 - Sistema esperto.....	34
Figura 4.1 - Fasi di un analisi Should Cost.....	52
Figura 4.2 - Combinazioni delle varianti delle caratteristiche principali del database.....	61
Figura 4.3 - Metodologia per lo sviluppo modelli di costo parametrici.....	64
Figura 5.1 - Inizializzazione del comando "Matrice di correlazione"	71
Figura 5.2 - Selezione del database	72

Figura 5.3 - Dati Tolleranza - % Errore - LD	78
Figura 5.4 - Influenza della presenza di geometrie fuori asse sulla precisione di LD.....	79
Figura 5.5 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)	80
Figura 5.6 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale - LD	81
Figura 5.7 - Andamento dell'errore percentuale in funzione del n° di utensili stimati - LD	82
Figura 5.8 - Costo medio per pezzi con n° utensili da 1 a 10 (A) e da 11 a 28 (B)	83
Figura 5.9 - Errore percentuale per le varie classi di produzione - LD.....	84
Figura 5.10 - Costo medio delle varie Classi.....	85
Figura 5.11 - Dati Tolleranza - % Errore - LC11	86
Figura 5.12 - Influenza della presenza di geometrie fuori asse sulla precisione di LC11.....	87
Figura 5.13 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B).....	87
Figura 5.14 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale – LC11.....	88
Figura 5.15 - Andamento dell'errore percentuale in funzione del n° di utensili stimati – LC11.....	89
Figura 5.16 - Costo medio per pezzi con n° utensili da 1 a 10 (A) e da 11 a 28 (B)	89
Figura 5.17 - Errore percentuale per le varie classi di produzione – LC11	90
Figura 5.18 - Costo medio delle varie Classi.....	91
Figura 5.19 - Dati Tolleranza - % Errore - LC9	92

Figura 5.20 - Influenza della presenza di geometrie fuori asse sulla precisione di LC9.....	92
Figura 5.21 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)	93
Figura 5.22 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale – LC9	94
Figura 5.23 - Andamento dell'errore percentuale in funzione del n° di utensili stimati – LC9.....	94
Figura 5.24 - Costo medio per pezzi con n° utensili da 1 a 10 (A) e da 11 a 28 (B)	95
Figura 5.25 - Errore percentuale per le varie classi di produzione – LC9	96
Figura 5.26 - Costo medio delle varie Classi.....	97
Figura 5.27 - Dati Tolleranza - % Errore – LC ACMA.....	98
Figura 5.28 - Influenza della presenza di geometrie fuori asse sulla precisione di LC ACMA.....	98
Figura 5.29 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)	99
Figura 5.30 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale – LC ACMA	100
Figura 5.31 - Errore percentuale per le varie classi di produzione – LC ACMA	100
Figura 5.32 - Costo medio delle varie Classi.....	101
Figura 6.1 - Numero casi a diverso MAPE - Tornitura	109
Figura 6.2 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 - Tornitura.....	110
Figura 6.3 - Numero casi a diverso MAPE - Tornitura v2	112

Figura 6.4 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 - Tornitura v2.....	113
Figura 6.5 - Numero casi a diverso MAPE – Fresatura.....	114
Figura 6.6 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 - Fresatura	115
Figura 6.7 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 – Fresatura v2	117
Figura 6.8 - Numero casi a diverso MAPE - Fresatura v2	117

IV. ELENCO DELLE TABELLE

Tabella 5.1 - Tabella dei costi unitari per diversi range di quantità.....	66
Tabella 5.2 - Tabella dei costi unitari a quantità unitaria (1) e quantità d'acquisto reale (n).....	66
Tabella 5.3 - Tabella riassuntiva dei costi stimati.....	66
Tabella 5.4 - Parametri prestazionali per analisi globale	68
Tabella 5.5 - Valori arbitrari per esempio indici.....	69
Tabella 5.6 - Matrice delle correlazioni completa.....	73
Tabella 5.7 - Matrice delle correlazioni ridotte.....	74
Tabella 5.8 - Parametri ad alta correlazione per la famiglia tecnologica della tornitura	75
Tabella 5.9 - Esempio di matrice delle correlazioni per la fresatura	76
Tabella 5.10 - Analisi tempi.....	102
Tabella 6.1 - MAPE Modello parametrico – Tornitura.....	109
Tabella 6.2 – R ² modello parametrico - Tornitura.....	109
Tabella 6.3 - Numero casi a diverso MAPE - Tornitura v2.....	112
Tabella 6.4 – R ² modello parametrico - Tornitura v2.....	112
Tabella 6.5 - MAPE Modello parametrico – Fresatura	114
Tabella 6.6 – R ² modello parametrico - Fresatura	114
Tabella 6.7 - MAPE Modello parametrico – Fresatura v2.....	116
Tabella 6.8 - R ² modello parametrico – Fresatura v2.....	117
Tabella 6.9 - Analisi qualità modello parametrico – Torniti.....	119
Tabella 6.10 - Analisi qualità modello parametrico – Fresatura	119
Tabella 7.1 - Parametri ad alta correlazione con % errore	122
Tabella 7.2 - Risultati modello parametrico – Tornitura.....	123
Tabella 7.3 - Risultati modello parametrico – Tornitura v2.....	124
Tabella 7.4 - Risultati modello parametrico - Fresatura.....	124

Tabella 7.5 - Risultati modello parametrico - Fresatura v2..... 125

0. ACMA

0.1. GENERALITÀ

A.C.M.A. (Azionaria/anonima Costruzioni Macchine Automatiche) è un'azienda italiana produttrice di macchine automatiche per il packaging primario e secondario. L'azienda fu fondata nel 1924 a Bologna ed ha diversi stabilimenti produttivi in Italia.

Le macchine sviluppate da ACMA sono rivolte principalmente ai settori di:

- Dolciumi (cioccolato, prodotti da forno, caramelle e gomme)
- Detersivi e Saponi (liquidi, polveri e barrette)
- Tè e caffè

ACMA, fa parte del gruppo COESIA. Quest'ultima consiste in un gruppo di società addette a sviluppare soluzioni industriali e di imballaggio basate sull'innovazione.

0.2. STORIA

L'anno di fondazione di ACMA risale al 1924. Gaetano Barbieri, aveva come scopo quello di meccanizzare il processo di imballaggio della polvere di idrolitina [9], così ACMA inizia la sua produzione di macchine automatiche.

Nel 1930 Bruto Carpigiani, primo di un'intera generazione di progettisti nel settore imballaggio, inventò la "Ruota Z", uno strumento che trasformava il movimento rotativo continuo in un movimento intermittente, rettilineo. Grazie a questa innovazione, si introdussero nuovi modelli come l'ACMA 722 che serviva a incartare le caramelle [11].

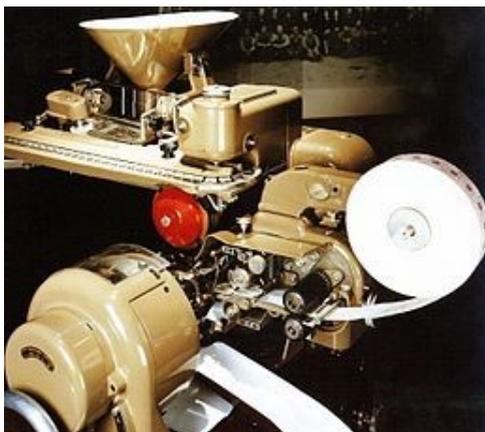


Figura 0.1 - Macchina automatica ACMA

Negli stessi anni '30 ACMA organizzò una rete di vendita che copriva i principali stati europei. Alla fine degli anni '40, riuscirono a produrre diverse macchine per il confezionamento di cibi e prodotti chimici: tra queste va annoverato il modello ACMA 721/1, la prima macchina che incartava saponette. Nel 1956 ACMA progettò la sua prima macchina per riempire scatole e cartoni di detergenti in polvere affermandosi all'estero. Nel 1963, con la scomparsa del fondatore Gaetano Barbieri, ACMA viene acquistata dall'Emhart Corporation. Tra gli anni '60 e '70 cominciò a farsi evidente la necessità di avere macchine per settori specifici, quindi il settore principale dell'azienda divenne il confezionamento, inscatolamento, strumenti per l'impacchettamento di saponi, detergenti, biscotti, lieviti. Nello stesso periodo ACMA rinnovò anche la tecnologia di produzione come il controllo numerico delle macchine².

Nel 1986 ACMA viene acquistata dalla CSII, in seguito rinominata Coesia, di cui ancora oggi fa parte. Nel 2012 ACMA viene integrata con Volpak, azienda spagnola specializzata in flexible pouching, ma a seguito della creazione della divisione operativa Consumer Goods Machinery di Coesia nel 2013, nel luglio 2015 le due società ritornano ad essere due entità separate sotto il controllo organizzativo della divisione, insieme ad altre 4 società del gruppo: GDM, Citus Kalix, Norden ed R.A Jones.

Ad oggi ACMA:

- Conta più di 250 dipendenti
- Reinveste circa il 5% del fatturato in R&D
- Ha prodotto più di 1000 macchine automatiche in 10 anni (di cui il 95% esportato all'estero).

0.3. DATI ECONOMICI

Dai dati finanziari consultabili in rete l'ultimo bilancio depositato da Azionaria Costruzioni Macchine Automatiche nel registro delle imprese corrisponde all'anno 2019 e riporta un range di fatturato di circa 60 milioni di euro [9] Si può dunque affermare che siamo di fronte ad un marchio leader del settore manifatturiero degli imballaggi e di una vera e propria eccellenza italiana conosciuta in tutto il mondo.



Figura 0.2 - Logo ACMA S.p.A.

1. INTRODUZIONE

Lo scopo cardine di ciascuna azienda è quello di massimizzare l'utile. Per fare ciò bisogna aumentare quanto più possibile la differenza tra costi e ricavi. I costi devono essere pienamente compresi, controllati e ridotti sin dalle prime fasi dello sviluppo del prodotto. A tal proposito negli ultimi anni si è cercato di individuare diverse soluzioni tecnico/funzionali finalizzate alla stima e riduzione dei costi già in fase di progettazione.

Il lavoro di tesi è stato svolto in collaborazione con ACMA S.p.A. ed ha come colonna portante l'area del "cost engineering" o, più in dettaglio l'analisi di metodologie di preventivazione del costo di componenti meccaniche progettate per le macchine automatiche prodotte da ACMA. In particolare, l'azienda è interessata ad indagare l'efficienza di alcuni software utilizzati internamente per la costificazione rispetto alle diverse famiglie tecnologiche, nonché ad aprire il processo di preventivazione dei costi a strumenti innovativi come modelli parametrici basati sul Machine Learning. Nel lavoro di tesi ci si è interessati anche alle tempistiche di esecuzione della preventivazione dei costi durante la fase progettuale per capire se sia conveniente fornire gli strumenti usati direttamente al progettista. Sebbene tale fase non sia di interesse specifico di questo ruolo, in un'ottica di una struttura aziendale simultanea, risulta sempre più utile rendere le fasi di sviluppo del prodotto il più contemporanee possibili per ridurre al minimo il time to market ed aumentare la competitività sul mercato. Bisogna qui precisare che ACMA si occupa della sola fase di progettazione del componente ma non della fase realizzativa, la quale viene delegata a fornitori terzi. L'ufficio acquisti sarà dunque chiamato a contrattare con i diversi fornitori per ottenere il miglior prezzo di mercato ed aumentare l'utile.

L'indagine che si va a fare si basa su due fattori principali:

1. Precisione del software
2. Tempo di utilizzo il software

Per "precisione del software" si intende quanto accuratamente il programma riesce a stimare il costo dei vari componenti. Per fare ciò si determinano, come si vedrà in seguito diversi insiemi di dati (torniti, fresati) in modo da dividere in famiglie omogenee i codici di partenza e trarre delle conclusioni su ciascuna di esse (per ciascun software e ciascun database).

Per quanto riguarda invece il "tempo di utilizzo del software" si va a studiare se in termini di ore/lavoro questa operazione di costificazione possa andare a ridurre l'efficienza del progettista in fase di disegno del componente. Questo dato risulta molto importante e viene considerato un parametro prestazionale di cui tener conto. Questo perché si può progettare una macchina efficientissima ad un ottimo prezzo, ma se per fare ciò si aumenta il tempo di progettazione di due/tre mesi, tutto il guadagno ottenuto viene a dissiparsi.

Un ulteriore passo che viene svolto in questa tesi consiste nel tentativo di andare a "costruire" uno strumento di preventivazione veloce e di facile utilizzo per i progettisti. Ci si propone, a tale scopo, di sfruttare algoritmi di machine learning per trovare un modello parametrico in grado di trasformare degli input di tipica conoscenza del progettista in un output di prezzo del prodotto.

In questo modo si riuscirebbe a ridurre al minimo il tempo che il progettista deve dedicare alla fase di costificazione e anzi, tramite le informazioni presenti sul CAD non si esclude come sviluppo futuro quello di creare un software, che legga tali informazioni e riesca a monitorare "in diretta" il costo approssimativo del componente disegnato.

Prima di fare tutto ciò ci sarà bisogno però di dare al lettore il background scientifico di cui ha bisogno per comprendere l'argomento (Capitolo 2), definire lo stato dell'arte fino ad oggi e spiegare le differenze di questo lavoro rispetto ai precedenti (Capitolo 3), definire la metodologia di lavoro utilizzata per affrontare la tesi (Capitolo 4), e sviluppare il caso studio (Capitoli 5 e 6) analizzandone i risultati ottenuti.

2. BACKGROUND SCIENTIFICO

2.1. DESIGN TO COST: PROGETTARE A COSTO TARGET

Il design to cost (DTC) è un approccio gestionale strutturato, sviluppato con l'obiettivo di stimolare tutti gli operatori aziendali ad intervenire sulle principali determinanti dei costi e della qualità, fin dalle primissime fasi della progettazione di un nuovo prodotto.

In altri termini, il DTC è uno strumento per attuare una politica strategica di target cost management (TCM).

Il target cost management si presenta come un approccio orientato alla pianificazione dei costi (cost planning) e al miglioramento dei costi (cost improvement), attraverso il rigoroso operare di team interfunzionali permanenti.

Il modello tradizionale prevede che le diverse funzioni aziendali (ricerca e sviluppo, marketing, progettazione, industrializzazione, acquisti, produzione, ecc.) vengano coinvolte nel processo una per volta e in serie. Di conseguenza coloro che progettano e disegnano il nuovo prodotto lavorano “in isolamento”, con una conoscenza alquanto ridotta delle attività svolte a valle e dell’impatto che le scelte di design e di progettazione possono determinare sulla producibilità del nuovo prodotto e sui relativi costi di produzione.

Il modello più efficace di progettazione del prodotto/processo richiede una concorrenzialità delle attività progettuali.

Il DTC diventa quindi il primo motore dell’approccio TCM in quanto si pone l’obiettivo di realizzare un prodotto con:

- Funzioni e prestazioni richieste dal cliente;
- Prezzo che il cliente è disposto a pagare;
- Posizionamento competitivo che l'azienda desidera conseguire;
- Livello di redditività che l'azienda desidera ottenere;
- Recuperi di efficienza da attuare lungo tutto il ciclo di vita del prodotto.

Il processo di DTC comincia dalle prime fasi di concezione (concept), ricerca e progettazione fino allo sviluppo e industrializzazione del prodotto.

Per comprendere quanto la fase progettuale abbia un impatto fondamentale sulla struttura di costo del prodotto, si deve considerare che, nonostante la maggior parte dei costi verrà materialmente sostenuta in sede di allestimento degli impianti, di predisposizione delle strutture e di gestione dell'attività produttiva, una percentuale compresa tra il 60% e l'80% del costo finale di un prodotto è di fatto irreversibilmente decisa dalle scelte effettuate in sede progettuale.

Se la qualità e la funzionalità di un prodotto devono essere conseguite attraverso opportune scelte in fase di progettazione, anche i livelli di costo (soprattutto questi) dipendono dalle soluzioni progettuali adottate. Quando il progetto è terminato anche se solo il 10-15% del costo totale è stato sostenuto, circa l'80% dei costi sono determinati (impegnati); in altre parole, benché l'85-90% dei costi sia ancora da sostenere, l'85-90% di questi sono stati già impegnati con le scelte di progettazione del prodotto e del processo, come si evince dalla Figura 2.1.

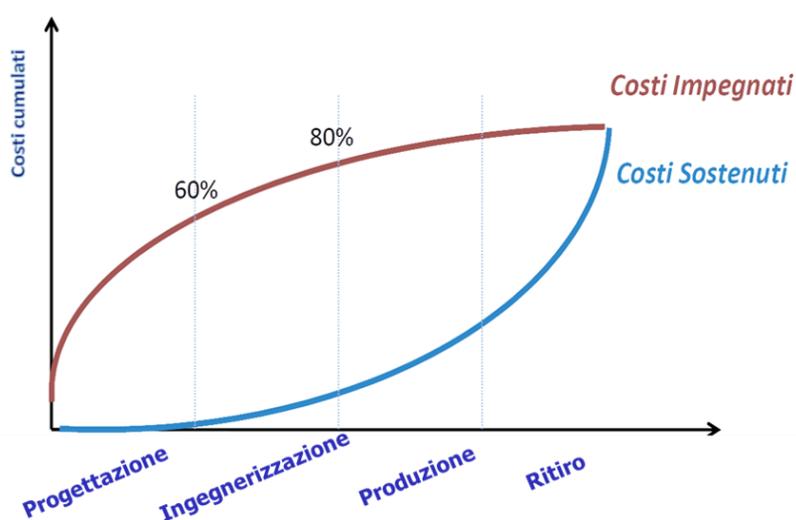


Figura 2.1 - Andamento dei costi sostenuti e impegnati durante il ciclo di vita del prodotto

Il costo delle modifiche di un progetto aumenta in modo inversamente proporzionale alle possibilità di ridurre il costo del prodotto. Risulta quindi necessario progettare il prodotto con obiettivi di costo chiari e definiti sin dalle prime fasi della progettazione.

Quanto detto collima con le logiche del cost management, secondo le quali già durante la fase di ideazione e di sviluppo di un prodotto devono essere definiti degli obiettivi economici di progettazione e di produzione del prodotto e dei suoi componenti.

Questi obiettivi devono riflettere sia le richieste dei clienti che le performance dei concorrenti: occorre spostare la sfida lanciata dal mercato all'interno dell'intera catena del valore.

2.2. TECNICHE PER LA STIMA DEI COSTI

In ogni fase dello sviluppo di un prodotto, l'azienda spende denaro e incorre in una serie di costi. Il costo di un prodotto o di un progetto è tanto un obiettivo in sé quanto lo è la prestazione tecnica del prodotto stesso.

Si può quindi dire che il costo è un elemento di design, ed è essenziale controllare questo parametro il più presto possibile per trovare il miglior compromesso in funzione del prezzo, per i progetti o il prodotto.

Tuttavia, la valutazione economica durante la fase di progettazione non è facile, ed è molto diversa dalla valutazione quando la progettazione del prodotto/processo è completa e dettagliata, caso in cui è possibile tenere conto del costo di tutte le scelte di ottimizzazione.

Purtroppo, nella fase di progettazione, il progetto o il prodotto non sono mai completamente definiti.

In questa fase è necessario implementare dei metodi di stima dei costi rapidi e più o meno precisi (in base ai dati disponibili), che consentano al progettista di selezionare una soluzione preferibile rispetto a un'altra per motivi economici.

Esistono svariate metodologie per la stima dei costi, ognuna delle quali presenta dei vantaggi e degli svantaggi.

La suddivisione presentata da A. Niazi suddivide le metodologie in due categorie principali [1]:

1. Tecniche qualitative;
2. Tecniche quantitative.

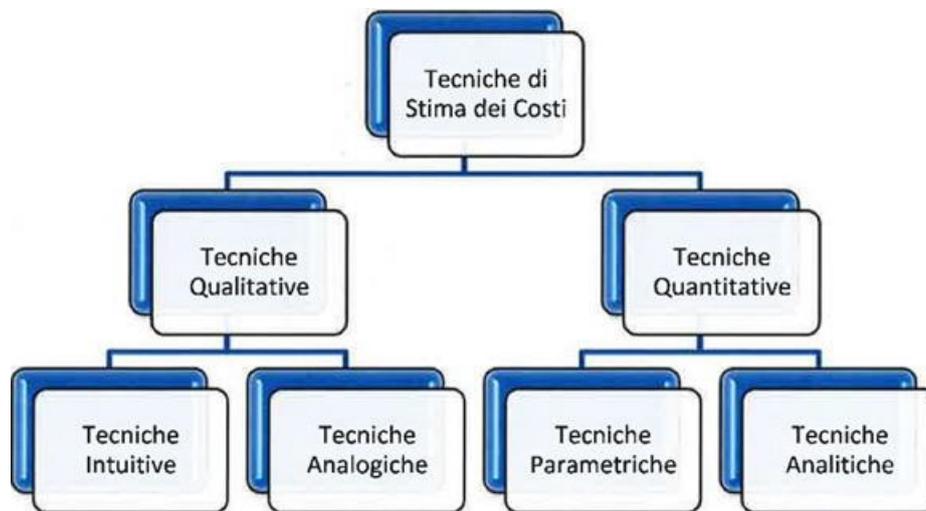


Figura 2.2 - Classificazione delle tecniche di stima dei costi

A loro volta, sulla base di questa suddivisione esistono tecniche con un approccio di tipo intuitivo, analogico, parametrico e analitico. Tutte le tecniche sono illustrate nel diagramma ad albero di Figura 2.2.

Le tecniche qualitative sono basate principalmente sull'esperienza e prevedono il confronto del nuovo prodotto con prodotti realizzati in precedenza al fine di riconoscerne le somiglianze. Per individuare le similarità possono essere utilizzati dei database di conoscenza e dei sistemi relazionali in cui vengono raccolte le caratteristiche distintive di un pezzo, ciò permette un più facile recupero delle informazioni.

In genere queste metodologie permettono di ottenere una stima di massima dei costi durante le prime fasi di ideazione del prodotto (progettazione concettuale). Le tecniche quantitative invece, sono basate su un'analisi dettagliata del progetto del prodotto, delle sue caratteristiche e dei processi produttivi.

Il costo totale può essere calcolato con una funzione di costo che tiene conto di diversi parametri, o come sommatoria di singole unità elementari che rappresentano le singole risorse coinvolte durante l'intero ciclo di produzione.

Sebbene queste tecniche siano note per fornire risultati più accurati, il loro utilizzo è normalmente limitato alle fasi finali del ciclo di progettazione a causa del requisito di una progettazione dettagliata del prodotto.

Detto ciò, resta chiaro che questi metodi non possono essere utilizzati durante tutto il ciclo di vita; alcuni metodi sono meglio di altri a seconda del contesto.

Nei seguenti paragrafi si fornisce un dettaglio maggiore delle tecniche qualitative e quantitative le quali includono anche lo strumento LeanCOST utilizzato nell'attività di costificazione eseguita per lo studio proposto.

2.2.1. Tecniche intuitive

Le tecniche intuitive sono basate sull'utilizzo dell'esperienza passata. La conoscenza può essere contenuta nelle persone o essere raccolta in modo più sistematico in sistemi più strutturati (ad esempio un tipo di software). La raccolta può essere fatta in vari modi, si possono utilizzare regole, alberi delle decisioni, giudizi, ecc.

Il database aiuta l'utente nel processo decisionale e nella stima del costo basandosi sulle informazioni in input relative ai parametri caratteristici dell'oggetto da realizzare.

L'esperienza di una persona esperta, o di un team di persone esperte è il metodo più comune attualmente utilizzato nelle aziende manifatturiere, queste effettuano un'analisi del componente e stimano il costo in base alla loro esperienza personale.

I vantaggi delle tecniche intuitive sono:

- Approccio rapido;
- Utilizzabili quando si hanno a disposizione poche informazioni;
- Facili da utilizzare

Gli svantaggi sono:

- Metodo soggettivo, la stima può variare di molto a seconda dell'esperienza personale;
- Richiede la disponibilità di persone esperte;
- Non sempre sono accurate e rigorose.

2.2.2. Tecniche analogiche

Queste tecniche fanno riferimento a dati storici sul costo di componenti prodotti in passato, stabilendo una relazione tra determinate variabili ritenute influenti e il costo ad esse associato; queste relazioni sono utilizzate per stabilire il costo di un nuovo componente.

I vantaggi delle tecniche analogiche:

- Facilmente implementabili quando si hanno dati disponibili;
- Accuratezza accettabile per pezzi semplici.

Gli svantaggi sono:

- Richiede la disponibilità di dati su prodotti simili;
- Richiede un affidabile comparazione con i prodotti esistenti;
- Poco accurate e difficilmente implementabili per oggetti molto diversi;
- Richiede un investimento significativo, non sempre compatibile con la dimensione delle imprese meccaniche che sono spesso medio – piccole.

2.2.3. Tecniche Analitiche

Queste tecniche richiedono la scomposizione del prodotto in unità elementari, operazioni e attività che rappresentano le diverse risorse utilizzate durante il ciclo produttivo; il costo totale del prodotto è espresso come la sommatoria delle singole componenti.

I vantaggi delle tecniche analitiche sono:

- Forniscono una stima molto accurata;
- Forniscono associata alla stima dei costi anche l'intero ciclo di lavorazione.

Gli svantaggi sono:

- Richiedono informazioni dettagliate;
- Richiedono molto tempo per essere sviluppate.

2.2.4. Tecniche Parametriche

L'utilizzo di tecniche parametriche prevede l'identificazione dei parametri principali del processo produttivo e tramite l'utilizzo di formule standard viene calcolato il costo del prodotto.

Nella maggior parte dei casi non è sufficiente utilizzare una formula lineare basata su un singolo parametro, ma data l'elevata variabilità delle caratteristiche è necessario utilizzare formule che tengano conto di più fattori contemporaneamente.

Questa metodologia può essere utilizzata quando si è in grado di definire in maniera chiara quali sono i parametri di processo che hanno un'elevata influenza sul costo.

I vantaggi delle tecniche parametriche sono:

- Buona affidabilità anche per componenti e prodotti molto diversi;
- Non richiedono persone esperte nell'utilizzo;
- Sono disponibili software commerciali basati su questo approccio.

Gli svantaggi delle tecniche parametriche:

- Richiedono la disponibilità di dati sulle relazioni accurate e verificate tra parametri e costi;
- Richiede l'utilizzo di un modello adeguato.

In particolare, in questa tesi sono state utilizzate tecniche di stima parametrica.

2.3. TECNICHE DI MACHINE LEARNING

In questo capitolo si vogliono quindi descrivere la natura e le caratteristiche principali delle tecniche di machine learning, le quali verranno poi utilizzate per ottenere quelli che sono i modelli di costo parametrici.

Come definito nel paragrafo precedente, la tecnica utilizzata per la stima di costo è quella parametrica e, in particolare, ci sono due approcci compatibili riguardanti la stima dei costi:

- Approccio tradizionale;
- Approccio innovativo.

Nel primo caso il costo è stato espresso con delle funzioni analitiche lineari definite "Relazioni di stima di costo" (CER – Cost Estimation Relationship) e costruite per mezzo di applicazioni di metodologie statistiche.

Nel secondo caso il costo è stato stimato tramite tecniche di machine learning che possono essere reti neurali, come l'artificial neural network o il deep learning, oppure tecniche basate sugli alberi decisionali come random forest e gradient boosting.

Da questi metodi di predizione scaturiscono differenti modelli ed è utile metterli a confronto per un medesimo progetto per determinare quale di questi è più conveniente utilizzare; in tale modo si evidenziano sia le capacità predittive che le carenze di ciascun modello.

Le stime di costo sono per loro stessa natura approssimative in quanto non consentono di determinare un valore puntuale del costo.

È previsto che il valore individuato in questa fase per la valutazione della fattibilità economica possa fluttuare all'interno di un range più o meno esteso in funzione di quelle che sono le conoscenze specifiche del progetto in esame e dei riferimenti precedentemente accumulati nel corso del tempo.

Si vuole con questi sistemi valutare l'ordine di grandezza dei costi da affrontare: non essere in grado di svolgere questa valutazione significa andare incontro a gravi problemi nelle fasi di produzione di un progetto, quali la mancanza di fondi nel caso di sottostima, oppure, nel caso opposto, l'immobilizzo di capitali che potrebbero essere utilizzati per altri fini.

Inoltre, è evidente che una stima di costo errata non permette di soddisfare le aspettative iniziali di progetto. Pertanto, questi sistemi e i modelli che verranno applicati sono strumenti di estrema utilità per decidere se procedere in una direzione o in un'altra nella produzione di un progetto, o addirittura se sia conveniente realizzare il progetto.

2.3.1. Regressione lineare

Il metodo principale con cui viene sviluppata la stima dei costi parametrici è la regressione. Questo metodo tenta di stabilire la natura della relazione tra le variabili fornendo un meccanismo di previsione.

Ci sono due diversi tipi di variabili: dipendenti designate dal simbolo y e indipendenti designate dal simbolo x . La variabile y rappresenta un tipo di costo, mentre le variabili x rappresentano vari parametri del sistema.

La regressione, quindi, è un ramo delle statistiche applicate che permette di quantificare la relazione tra la variabile dipendente e una o più variabili indipendenti e di descrivere l'accuratezza di tale relazione.

Sebbene possa essere non lineare, sono state analizzate sole le forme lineari di regressione (semplice e multipla).

Anche se rimanere nel campo della linearità può sembrare una semplificazione eccessiva del problema, ci sono diverse buone ragioni per adottare questo approccio. I costi dovrebbero logicamente, e spesso lo fanno, variare in modo lineare con la maggior parte delle caratteristiche fisiche e prestazionali.

Inoltre, molte funzioni curvilinee ed esponenziali possono essere trasformate in una forma lineare, prestandosi così all'analisi lineare.

2.3.2. Regressione lineare semplice

Si parla di regressione lineare semplice quando è necessaria una sola variabile indipendente per stimare il valore di y . La variabile y è correlata alla variabile x dalla seguente espressione:

$$y = \beta_0 + \beta_1 x \quad (1)$$

Dove:

β_0 = Intercetta

β_1 = Coefficiente angolare

x = Variabile indipendente

y = Variabile dipendente

I coefficienti β_1 e β_2 possono essere calcolati graficamente tramite interpolazione (Figura 2.3). [3]

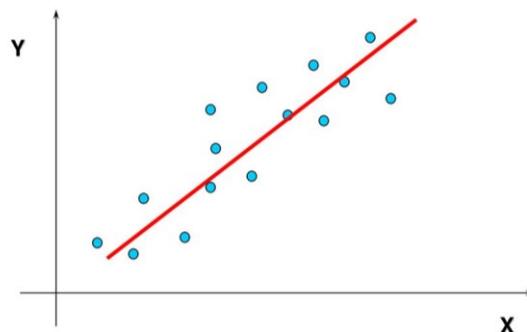


Figura 2.3 - Regressione lineare semplice

2.3.3. Regressione lineare multipla

A differenza di quella semplice è caratterizzata da n variabili dipendenti.

L'equazione assumerà la seguente forma:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n \quad (2.1)$$

Dove

β_0 = Intercetta

β_1 = Coefficiente angolare di x_1

β_2 = Coefficiente angolare di x_2

x = Variabile indipendente

y = Variabile dipendente

I coefficienti β_1 e β_2 possono essere calcolati matematicamente.

Sia per la regressione lineare semplice che per la regressione lineare multipla, il coefficiente di determinazione R^2 permette la misura dell'accuratezza delle equazioni. Può avere un valore compreso tra -1 e +1. In Figura 2.4 si vede chiaramente che maggiore è il numero, migliore è l'adattamento della linea di regressione ai dati effettivi.

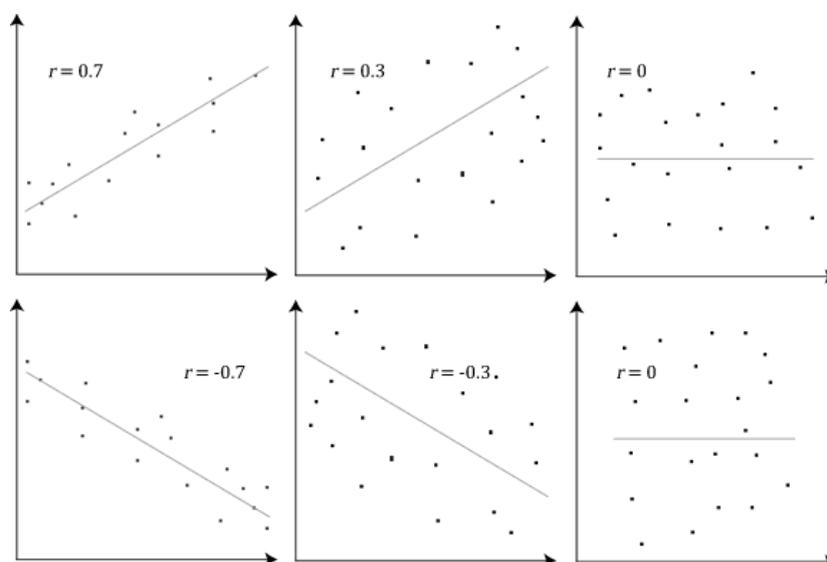


Figura 2.4 - Andamento coefficiente R^2

2.3.4. Artificial Neural Network

Le reti neurali, definite dall'acronimo ANN (Artificial Neural Networks) sono un approccio innovativo di machine learning che permettono di sviluppare problemi computazionali complessi in maniera veloce.

Esse si ispirano alle funzionalità del cervello umano ed alla sua struttura, che può essere rappresentata come una rete di neuroni interconnessi tra loro e in grado di accumulare nel tempo le conoscenze in maniera "distribuita": l'informazione viene codificata per mezzo di impulsi elettrici nei neuroni ed è immagazzinata modificando la struttura molecolare e fisica delle connessioni [4].

Analogamente, l'ANN è costituita da tante "unità" chiamate neuroni che, considerando l'architettura di rete più comune ("Multilayer Perceptron") (Figura 2.5), sono disposti in strati successivi: ciascun neurone è tipicamente collegato a tutti i neuroni dello strato successivo tramite connessioni pesate o sinapsi. Una connessione è un valore numerico (il "peso" appunto). È possibile individuare tre strati principali:

- Strato di input (input layer): costituito da neuroni che rappresentano la sorgente dei dati;
- Strato di output (output layer): costituito da neuroni che producono la risposta finale della rete;
- Strato intermedio (hidden layer): può essere più di uno e contiene dei neuroni nascosti.

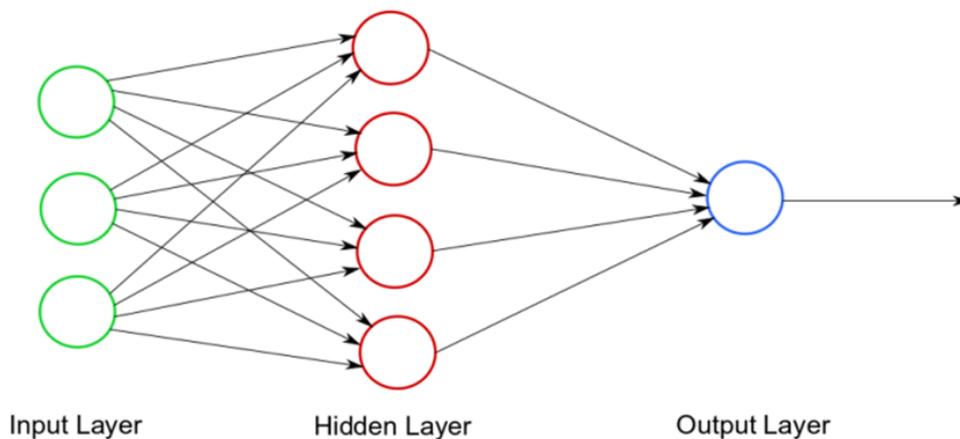


Figura 2.5 - Architettura reti neurali

Ciascun neurone dello strato intermedio, somma i valori pesati di tutti i neuroni ad esso collegati e aggiunge un valore di *bias*. A questo risultato viene applicata una funzione di attivazione che trasforma matematicamente il valore prima di passarlo allo strato successivo. In questo modo i valori di input vengono propagati attraverso la rete fino a restituire l'output.

Oltre alla struttura, un'altra similitudine fra la rete neurale ed il cervello umano è la capacità di apprendere.

Un ANN per funzionare correttamente ha bisogno di essere allenata, cioè si richiede all'utilizzatore di immagazzinare una sempre maggior quantità di dati dettagliati e di informazioni che possano addestrare la rete neurale (tali dati sono chiamati "patterns"). Questa operazione permette di determinare la struttura definitiva della rete e quindi di determinare i valori dei parametri caratteristici (pesi e bias).

Da quanto detto, una rete può essere vista come un sistema a scatola chiusa in grado di fornire un output a fronte di un dato input.

Sono tre gli aspetti principali che descrivono e caratterizzano una ANN:

1. Architettura della rete: definita dal numero di neuroni per ogni strato, il numero di strati intermedi e le connessioni tra i neuroni;

2. Tipologia della funzione di attivazione: tipicamente non lineare (funzione sigmoideale o funzione tangente iperbolica);
3. Modalità di apprendimento: può essere di tipo supervisionato o non supervisionato. Nel primo caso i pesi vengono modificati sulla base dell'errore commesso dalla rete rispetto ai dati reali di output (ciò è possibile se si è a conoscenza di coppie di input-output). Nel caso del non supervisionato invece i pesi variano nel corso dell'apprendimento in base ad una regola definita a priori che non utilizza l'errore rispetto al dato del parametro dipendente.

2.3.4.1. Algoritmo di Backpropagation

Tra le diverse tecniche di apprendimento supervisionato la più utilizzata è l'algoritmo di Backpropagation, BP. L'obiettivo di questa tecnica è quello di determinare l'intensità delle connessioni tra i nodi e quindi i valori dei parametri caratteristici (pesi e bias).

Tale algoritmo è costituito da due fasi: forward (in avanti) e backward (all'indietro).

In particolare, nella fase forward i pattern di apprendimento (coppie di dati input-output) sono presentati ai nodi di input. La risposta della rete si sviluppa lungo i livelli intermedi fino ad arrivare ai nodi di output. Durante questa fase i pesi ed i bias rimangono invariati.

Nella fase backward, l'errore esistente tra il processo reale e la risposta della rete viene calcolato e propagato all'indietro attraverso i nodi. Tramite opportune formule di aggiornamento i pesi ed i bias vengono modificati fino al primo livello della rete (quello di input).

L'aggiornamento dei parametri caratteristici nel corso dell'addestramento può essere visto come un problema di ottimizzazione di una funzione obiettivo volta a minimizzare l'errore medio compiuto sull'insieme dei pattern di apprendimento. Un tipico problema di questa tecnica è il problema dell'overfitting (detto anche overlearning): si verifica quando la rete apprende in modo ottimo la risposta ai pattern di input-output perdendo però la capacità di generalizzazione e di risposta a dati di input non ancora sperimentati.

Il processo di apprendimento termina con una fase finale di convalida dove, tramite l'utilizzo di pattern diversi da quelli di apprendimento, si verifica se l'algoritmo fornisce risultati soddisfacenti.

2.3.4.2. Cross validation

La cross-validation (o convalida incrociata) è una tecnica statistica usata nel machine learning per eliminare il problema dell'overfitting nei training-set. In particolare, la convalida incrociata cosiddetta k-fold si sviluppa nel seguente modo:

1. Si dividono i dati di input in k sottoinsiemi di dati (noti anche come fold);
2. Si addestra il modello su tutti i sottoinsiemi tranne uno (k-1) e quindi si valuta il modello sul sottoinsieme che non è stato utilizzato per l'addestramento. Ciò significa che ogni volta uno dei sottoinsiemi k viene usato come set di test, mentre gli altri sottoinsiemi k-1 sono messi insieme per formare un set di allenamento (Figura 2.6) [5];

Il processo viene ripetuto k volte, ogni volta con un diverso sottoinsieme riservato per la valutazione (ed escluso dall'addestramento).

L'esecuzione di una k-fold genera k modelli di machine learning, k origini dati di training e k origini dati di testing. Per ogni modello viene generato un parametro delle prestazioni e facendo la media si ottiene la prestazione complessiva.

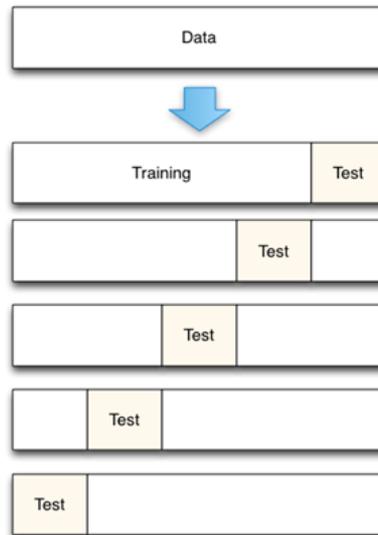


Figura 2.6 - Cross validation

In particolare, si vedrà in seguito che il concetto del cross validation sarà applicato per la creazione e l'allenamento del modello di costo di ogni singola tecnica.

2.4. DEEP LEARNING

Il deep learning è una particolare rete neurale che a differenza della rete neurale semplice, risulta essere composta dal almeno due livelli nascosti. Tipicamente le applicazioni Deep Learning tendono ad avere molti strati nascosti, nell'ordine della decina di unità (Figura 2.7).

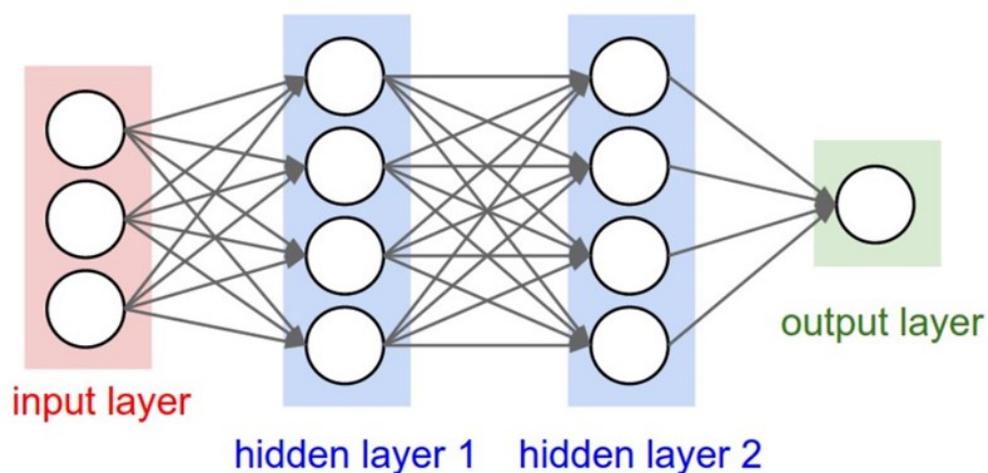


Figura 2.7 - Struttura Deep Learning

Si può immaginare il deep learning come un sistema che sfrutta una classe di algoritmi di apprendimento automatico, dove ciascun livello successivo utilizza l'uscita del livello precedente come input.

Gli algoritmi possono essere sia di tipo supervisionato sia non supervisionato e le applicazioni includono l'analisi di pattern (apprendimento non supervisionato) e classificazione (apprendimento supervisionato).

Inoltre, le caratteristiche di più alto livello vengono derivate da quelle di livello più basso per creare una rappresentazione gerarchica.

Di fatto applicando il deep learning, si ha quindi una macchina che riesce autonomamente a classificare i dati ed a strutturarli gerarchicamente, trovando quelli più rilevanti e utili alla risoluzione di un problema, migliorando le proprie prestazioni con l'apprendimento continuo.

È possibile dire che risulta essere la tecnica che più ripercorre le connessioni biologiche del cervello umano, le quali potenzialità trovano spazio in qualsiasi applicazione.

2.4.1. Random Forest

Il Random Forest è un algoritmo di machine learning molto popolare per la sua semplicità, facilità d'uso e interpretabilità; è una metodologia di apprendimento supervisionato per ensemble che appartiene alla famiglia di algoritmi dell'albero delle decisioni.

Un albero delle decisioni rappresenta un modello di classificazione o regressione in una struttura ad albero. Ogni nodo nella struttura ad albero rappresenta una particolare "domanda" su una caratteristica (feature), mentre ogni ramo rappresenta una decisione e ogni foglia alla fine di un ramo rappresenta il valore di output corrispondente (Figura 2.8). [3]

Per ottenere un risultato, partendo da uno specifico input, il processo decisionale inizia dal nodo radice (in cima) e percorre l'albero fino a raggiungere una foglia che contiene il risultato. In ogni nodo, il percorso da seguire dipende dai valori assunti dalle varie feature. Simile alle reti neurali, l'albero viene creato tramite un processo di apprendimento usando i dati di training.

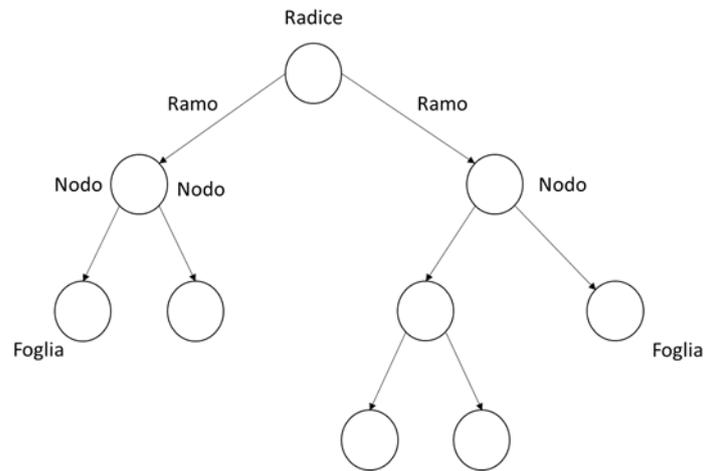


Figura 2.8 - Albero delle decisioni

Il Random Forest è costituito da un insieme di alberi decisionali diversi tra loro ma con la stessa origine di dati di training. Ogni albero delle decisioni viene creato da un sottoinsieme diverso e casuale dei dati di training.

Questa suddivisione permette di risolvere un tipico problema: se si creasse un unico albero decisionale per l'intero insieme di dati di training, si andrebbe incontro a un modello di predizione con una scarsa affidabilità dovuta alla presenza di overfitting sul training-set e di una varianza elevata.

Invece, il Random Forest è costituito da n alberi deboli ma che nel complesso costruiscono un insieme di "decision maker" in cui il risultato collettivo può essere creato con il voto a maggioranza in caso di classificazione o media in caso di regressione (Figura 2.9). [3]

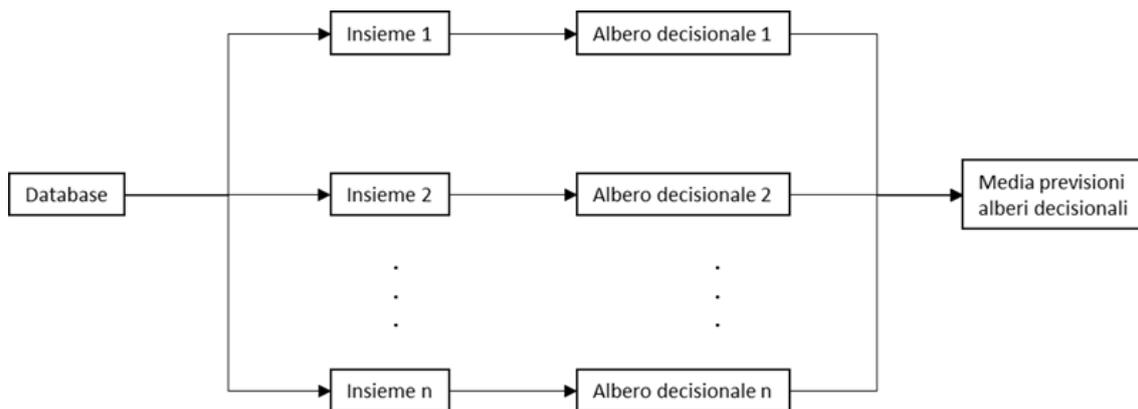


Figura 2.9 - Struttura Random Forest

Ciò compensa i potenziali errori dei singoli alberi in modo che si riduca la varianza e il modello sia più preciso.

2.4.2. Gradient Boosting

Il principio base di un algoritmo Gradient Boosting è che è un modello esemble, ossia si basa su una combinazione di singoli modelli semplici, detti weak learner, che insieme creano un nuovo modello più potente chiamato strong learner.

Il modo in cui questa combinazione viene fatta è appunto definito tecnico boosting, in quanto inizia inserendo un modello iniziale, come ad esempio un albero o una regressione lineare, nei dati.

In seguito, viene creato un secondo modello in sequenza, che si concentra sulla previsione accurata dei casi in cui il primo modello ha prestazioni scarse; la combinazione di questi due modelli dovrebbe essere migliore di entrambi i modelli presi singolarmente.

In linea generale si può dire che il gradient boosting ragiona contrariamente rispetto alle tecniche descritte fino ad ora; infatti, anziché puntare a prevedere il risultato della variabile che vogliamo prevedere, si prefigge di prevedere gli errori del modello. Gli errori vengono stimati dai residui, ossia dalla differenza tra il valore realmente osservato e il valore previsto dal modello.

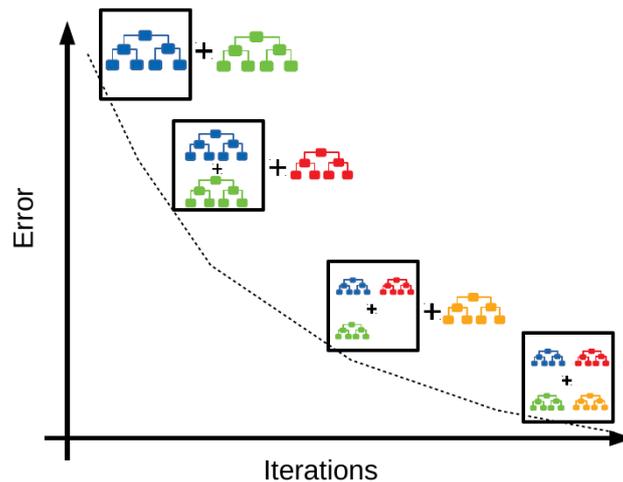


Figura 2.10 - Schema apprendimento Gradient Boosting

I residui vengono calcolati da alberi decisionali a più livelli di profondità e una volta calcolati i residui, il gradient boosting si avvale di un altro modello (solitamente sempre un albero di regressione) per prevedere i nuovi residui; iterando il modello punta a stimare i residui che, man mano, si riducono progressivamente.

Alla conclusione dell'algoritmo, l'errore complessivo del modello è ridotto rispetto all'inizio (Figura 2.10) [6].

Particolare interesse lo giocano i parametri caratteristici della tecnica stessa, ovvero la scelta del numero di alberi da utilizzare e il tasso di apprendimento.

Il primo rappresenta il numero di alberi che saranno adattati in serie per correggere gli errori di predizione, mentre il tasso di apprendimento, anche detto Learning Rate (LR), corrisponde alla velocità con cui l'errore viene corretto da ciascun albero al successivo ed è un semplice moltiplicatore che ricade nell'intervallo (0,1] con 0 escluso e 1 incluso.

Vari studi consigliano di impostare un LR basso piuttosto che alto in modo da ridurre la varianza complessiva del modello finale. [6]

Questo tipo di tecnica si presta molto bene a problemi di regressione, come si vedrà poi nel caso studio in esame.

2.5. STUDIO DI FABBRICAZIONE

Si definisce studio di fabbricazione l'insieme delle attività tese a trasformare il progetto di dettaglio in progetto di fabbricazione. Il progetto di dettaglio redatto dal progettista individua le caratteristiche di funzionamento e le prestazioni richieste dal prodotto, il progetto di fabbricazione definisce invece le caratteristiche e modalità di realizzazione del prodotto.

Lo studio di fabbricazione permette di generare il piano di processo per la realizzazione di uno specifico prodotto, ovvero:

- ciclo di fabbricazione per produrre le singole parti
- cicli delle operazioni per assemblare i vari componenti e/o sotto assemblati per ottenere il prodotto finito.

Possiamo quindi dire che il piano di processo è la rappresentazione univoca della modalità di realizzazione di una sola tipologia di prodotto. Si è messo in evidenza come lo studio di fabbricazione non serva solo alla definizione delle operazioni necessarie a realizzare il prodotto, ma contribuisce anche alla definizione del suo costo, alla valutazione delle inefficienze di produzione, ed alla scelta / progettazione del sistema di produzione.

Per quanto riguarda la definizione del costo, il piano di processo rappresenta la base per preparare il preventivo per l'acquisizione dell'ordine o per definire il prezzo di vendita del prodotto. Si è evidenziato come errori nella valutazione del costo possono portare al mancato ottenimento dell'ordine o a una riduzione dei ricavi.

Uno studio di fabbricazione può risultare complesso o meno. L'aumentare della complessità del piano cresce con il numero di processi necessari e con il numero di componenti che costituiscono il prodotto finale. Una parte complessa, infatti, può non essere riproducibile attraverso un unico processo e un'unica macchina. Spesso anche la sequenza delle operazioni non è univoca.

Gli input di uno studio di fabbricazione sono:

- disegni funzionali di componenti, sotto assemblati e prodotto finito;
- distinta dei materiali;
- quantità da produrre;
- tempi di consegna;

in output avremo:

- ciclo di fabbricazione e/o assemblaggio;
- costo del prodotto.

2.5.1. Metodi per effettuare uno studio di fabbricazione

I metodi per effettuare lo studio di fabbricazione non sono univoci. Si possono utilizzare due approcci:

- Manuale;
- Assistito dal calcolatore.

Utilizzando metodi manuali le scelte saranno legate all'esperienza e all'abilità del pianificatore di processo e si potranno avere inconvenienti quali la difficoltà nell'ottimizzazione del piano di processo e un incremento dei costi rispetto a quelli ottimali.

Se invece si è assistiti da un calcolatore (sistemi CAPP) si possono superare tali inconvenienti ed ottenere un piano ottimale.

Indipendentemente dal metodo utilizzato si possono adottare due tipi di approccio:

- Sequenziale: la pianificazione viene fatta prima della realizzazione
- Simultaneo: la realizzazione avviene contemporaneamente alla progettazione e tutte le altre attività.

A valle dell'attività di progettazione si ha l'attività di pianificazione dei processi che riceve come input gli output della progettazione. Bisognerà quindi determinare le parti da produrre internamente e quelle da acquistare da terzi. Per le parti da acquistare dunque la pianificazione è eseguita dal fornitore talvolta in collaborazione con il committente. Una delle attività più importanti della pianificazione consiste nella definizione dei parametri di processo, ad esempio, andare a definire le velocità di taglio, di avanzamento, la profondità di passata per le lavorazioni per asportazione di truciolo, così come la lavorazione e la macchina più adatta da utilizzare.

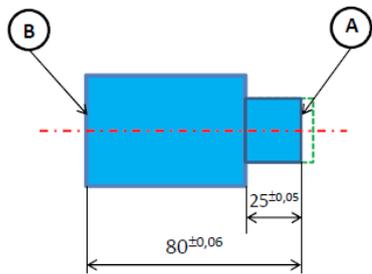
2.5.2. Pianificazione di processo

Il piano di processo è l'insieme di operazioni necessarie per fabbricare una singola parte attraverso una successione di processi tecnologici per ottenere il prodotto finito. La generazione del piano è basata sulla conoscenza di tutti gli aspetti relativi alla fabbricazione o all'assemblaggio del prodotto da realizzare. La pianificazione di processo rappresenta l'anello di congiunzione tra progettazione e produzione. La qualità la progetta il progettista, ma sarà il tecnologo a conferire realmente le caratteristiche richieste al pezzo. Si può pianificare il processo attraverso un approccio sequenziale o simultaneo alla progettazione. Questa fase ha più soluzioni: la sequenza di processo non è unica ma può avere diverse soluzioni. Non esiste inoltre un modello giusto da seguire per ottenere un piano di processo ottimale. Le scelte sono dovute all'esperienza, alle informazioni disponibili all'interno dell'azienda ed alle prescrizioni dei costruttori di macchine utensili.

Gli elementi fondamentali del piano di processo sono:

1. Operazione;
2. Sottofase;
3. Fase.

1. L'operazione è il più piccolo elemento in cui possiamo suddividere le attività di fabbricazione o di assemblaggio di una parte o di un prodotto. L'operazione non è ulteriormente suddivisibile. Essa viene realizzata su una specifica macchina, con un posizionamento specifico, con un utensile specifico.

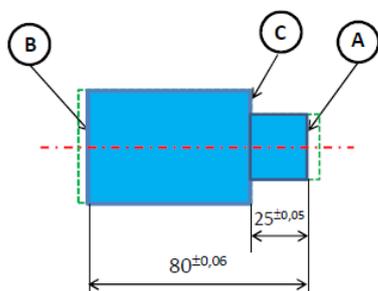


Linea tratteggiata: grezzo
Linea continua: parte finita

- Pezzo cilindrico sottoposto a tornitura frontale per la lavorazione della superficie A
- Posizionamento effettuato bloccando il pezzo sulla superficie B
- Operazioni:
 - sgrossatura (1^a operazione)
 - finitura (2^a operazione)

Figura 2.11 - L'operazione

2. Sottofase: è costituita dalle operazioni realizzate con uno specifico posizionamento del pezzo in una macchina. A cambiare del posizionamento del pezzo sulla macchina si definisce una nuova sottofase.



Linea tratteggiata: grezzo
Linea continua: parte finita

- Pezzo sottoposto a tornitura frontale per la lavorazione delle due superfici di base A e B
- Necessità di effettuare due posizionamenti del pezzo nella piattaforma autocentrata del tornio ⇒ presenza di due sottofasi:
 - 1° bloccaggio per lavorare la superficie A
 - 2° bloccaggio per lavorare la superficie B

Figura 2.12 - La sottofase

3. Fase: insieme delle sottofasi realizzate nella medesima macchina. Se cambio macchina si definisce una nuova fase.

Un pezzo richiede tante fasi quante sono le macchine necessarie alla sua produzione. Se una fase richiede più posizionamenti avrà per ogni fase “n” sottofasi. Se ogni sottofase necessita di “m” utensili avrò “n x m” operazioni.

Piano di processo	Fase 10	Sottofase 10.1	Operazione 10.1.1 Operazione 10.1.2 Operazione 10.1.3
		Sottofase 10.2	Operazione 10.2.1
		Sottofase 10.3	Operazione 10.3.1 Operazione 10.3.2
	Fase 20	Sottofase 20.1	Operazione 20.1.1 Operazione 20.1.2
		Sottofase 20.2	Operazione 20.2.1
		Sottofase 20.3	Operazione 20.3.1
	Fase 30	Sottofase 30.1	Operazione 30.1.1
		Sottofase 30.2	Operazione 30.2.1
		Sottofase 30.3	Operazione 30.3.1 Operazione 30.3.2
		Sottofase 30.4	Operazione 30.4.1

Figura 2.13 - Esempio di un piano di processo

2.5.3. Attività per la generazione del piano di processo

Le attività per la generazione del piano di processo sono:

- Analisi del disegno del prodotto;
- Analisi dei dati di partenza;
- Scelta delle fasi e sequenziamento;
- Scelta delle sottofasi e sequenziamento;
- Scelta degli utensili da utilizzare in ciascuna operazione;
- Scelta dei parametri di processo di ciascuna operazione;
- Scelta e progettazione delle attrezzature da utilizzare;
- Scelta degli strumenti e delle procedure di controllo;
- Calcolo dei tempi di fabbricazione e assemblaggio;
- Stesura del foglio di pianificazione.

2.5.4. Sistemi CAPP

Tutte le operazioni descritte nel paragrafo 2.5.3 possono essere svolte dai sistemi CAPP.

Essi possono essere generalmente di tre tipologie:

1. CAPP Varianti;
2. CAPP Generativi;
3. CAPP Semi-generativi.

Il CAPP variante dispone di un database contenente tantissimi dati di ciascuna famiglia di codici. Per famiglia di codici si intende un insieme di codici accomunati da caratteristiche geometriche e realizzative.

A ciascuna famiglia si associa un piano standard, ovvero l'insieme di tutte le lavorazioni, fasi e sottofasi necessarie per realizzare i pezzi appartenenti alla famiglia stessa.

Il piano di processo di una nuova parte viene successivamente ricavato attraverso l'associazione del nuovo codice alla famiglia di appartenenza, identificando quindi il piano di processo standard che dovrà essere modificato manualmente dal tecnologo.

La qualità del piano dipende molto dalle capacità del tecnologo in quanto il software è solo uno strumento di assistenza per la stesura del piano di processo finale.

Al contrario succede per i CAPP generativi, attraverso cui il programma riesce a generare un piano di processo senza riferimenti a piani preesistenti. Per far ciò il software ha bisogno di ricevere in input le informazioni sul modello CAD e di avere al proprio interno un database delle conoscenze affiancato ad una logica decisionale che lo possa guidare nelle giuste scelte.

La logica decisionale può essere di diversi tipi:

- Alberi delle decisioni: è rappresentato graficamente da un nodo radice (obiettivo), dai rami (contenenti le operazioni logiche) e da foglie (che rappresentano le possibili scelte se tutte le operazioni logiche che portano ad essa vengono rispettate);

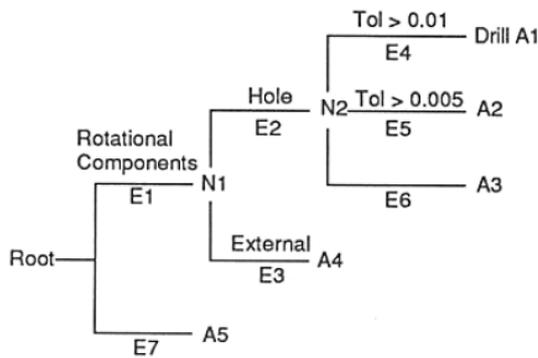


Figura 2.14 - Albero delle decisioni

- Tavole delle decisioni: la logica è simile a quella dell'albero delle decisioni ma la struttura è rappresentata mediante delle tabelle in cui sulle righe si hanno le condizioni e le azioni, mentre sulle colonne le regole decisionali;

Condizioni	C1	Y	Y	N	N
	C2	Y	N	Y	N
Azioni		A	B	C	D

Figura 2.15 - Tabella delle decisioni

- Sistemi esperti: sistemi che utilizzano l'intelligenza artificiale.

In genere un sistema esperto è costituito da tre moduli:

- Motore inferenziale: esegue le deduzioni e le derivazioni logiche sulla base di una serie di assunzioni e regole che devono essere applicate al problema;
- Meccanismo di spiegazione: spiega i risultati ottenuti da punto sopra;

- Interfaccia con l'utente: permette all'utente di sottoporre il problema al sistema esperto e di riceverne la soluzione.

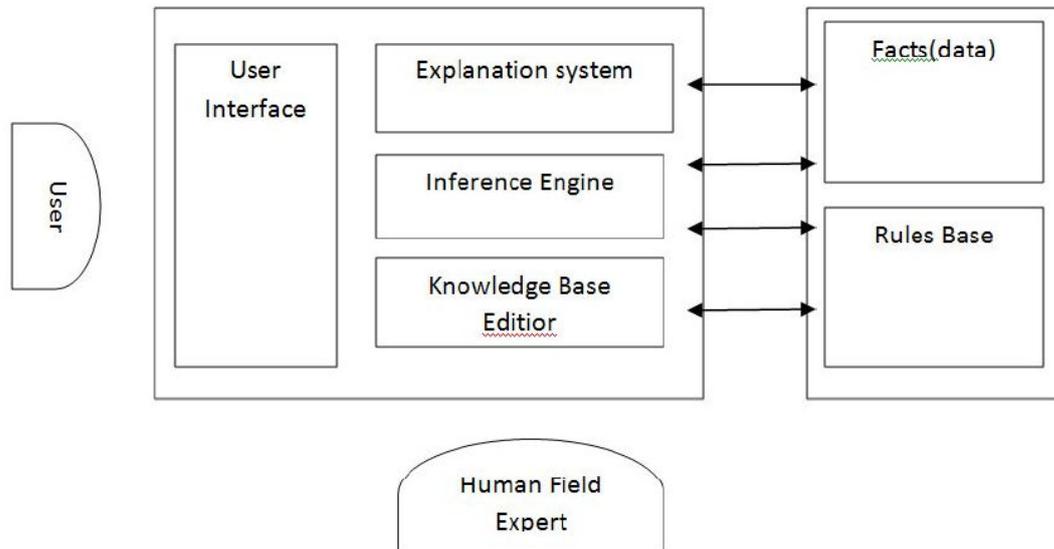


Figura 2.16 - Sistema esperto

Per quanto riguarda invece i sistemi CAPP semi-generativi possiamo dire che sono un misto tra i primi due casi appena descritti.

3. STATO DELL'ARTE

3.1. STRUMENTI PER PREVENTIVAZIONE DEI COSTI

In letteratura si possono trovare vari esempi di strumenti per la stima dei costi di produzione, sviluppati dalla fine dagli anni '70 fino ad oggi. Esistono numerosi strumenti commerciali e molte organizzazioni hanno sviluppato programmi proprietari per la preventivazione dei costi. La complessità di questi strumenti spazia dai fogli di calcolo a sistemi di database mainframe multiutente. I sistemi possono essere altamente specifici per un certo tipo di famiglia tecnologica o estremamente generici e utilizzabili per qualsiasi tipo di componente.

Una delle prime applicazioni può essere trovata da Orady et al. [20], che ha sviluppato uno strumento di stima assistita da computer per il calcolo dei tempi di produzione per componenti torniti. Il sistema calcola il tempo totale di lavorazione di un componente, considerando il tempo di set-up, il tempo di lavorazione e il tempo di carico/scarico macchina. In input l'utente inserisce dati come le dimensioni della macchina, dimensioni geometriche del pezzo, nome del processo produttivo ecc.

Siadat [21] invece utilizza un sistema di stima basato sull'entità di costo. Il sistema è in grado di calcolare il costo della lavorazione considerando le caratteristiche del pezzo.

Bouaziz [22] presenta un sistema di stima dei costi per la produzione di stampi basato sull'approccio analogico/analitico. Questo principio fa ricorso all'approccio analogico per la ricerca di analogie tra le forme da lavorare prima di raggrupparle in lavorazioni complesse. Per ogni parametro caratteristico il sistema genera un processo da utilizzare come campione e di conseguenza un modello di tempo di lavorazione. In una seconda fase viene utilizzato l'approccio analitico, con il quale si determina il tempo di taglio una volta scelta la velocità di asportazione per le operazioni di sgrossatura e/o finitura.

Ou-Yang [23] presenta un sistema che cerca di stimare il costo di fabbricazione di un disegno in base alle forme e alla precisione delle sue caratteristiche. L'obiettivo di questa ricerca è fornire uno strumento per aiutare un progettista con poca conoscenza del processo di produzione, a stimare il costo di fabbricazione durante la sua fase concettuale.

Jung [24] presenta un sistema di stima dei costi basato sulle caratteristiche per le parti lavorate. In questo modello il costo di lavorazione è proporzionale al tempo di lavorazione, che include il tempo operativo e il tempo non operativo.

Sajadfar [25] definisce un sistema di costificazione basato su tecniche di regressione e algoritmi di data mining per dati empirici riguardanti la saldatura.

Naranje [26] propone invece un sistema in grado di stimare il costo di produzione, il costo degli utensili e altri costi coinvolti nella produzione di parti imbutite.

Sono reperibili molti altri studi ma ci si vuole ora concentrare sulla situazione commerciale attuale descrivendo alcuni dei più noti software di preventivazione dei costi.

3.2. SOFTWARE SUL MERCATO

Attualmente, in commercio sono disponibili parecchi software per la stima dei costi. Esistono programmi di determinazione dei costi incorporati nei programmi CAD, come i sistemi PLM (Product Life Cycle Management), e programmi autonomi che possono lavorare anche separatamente da essi. [27].

3.2.1. Software di costo CAD-based

Uno dei software di determinazione dei costi CAD-based è SolidWorks Costing, [35], un modulo completamente integrato in SolidWorks.

Lo strumento aiuta i progettisti a prendere decisioni in base al costo di produzione e aiuta i produttori a creare preventivi per i clienti.

Il software crea stime automatiche dei costi di produzione per i vari processi utilizzando modelli incorporati e dati personalizzati.

Le informazioni sulla produzione e sui materiali nei modelli guidano lo strumento alla determinazione del costo di produzione di un determinato componente.

Nei modelli sono specificati il materiale utilizzato, i processi di produzione (come taglio laser, piegatura o fresatura), il metodo di produzione (lavorazione, fusione, plastica stampata, stampa 3D) e i costi associati ad essi.

SolidWorks Costing è utile per le diverse figure professionali:

- dal punto di vista del designer, fornisce stime di quanto dovrebbero costare le parti per la produzione;
- dal punto di vista dei produttori, crea preventivi accurati in base ai materiali, ai processi e ad altri costi associati necessari per la produzione delle parti.

Il tool crea un processo di preventivazione più veloce rispetto ai metodi manuali, aiuta a eliminare gli errori e fornisce un sistema di costificazione accurato e ripetibile che è possibile aggiornare ogni volta che i costi del materiale o della manodopera necessitano di revisione.

3.2.2. Software di costo generali

3.2.2.1. SEER for Manufacturing

SEER for Manufacturing è un programma autonomo per la preventivazione dei costi di componenti meccaniche [29].

SEER-MFG è uno strumento progettato per consentire ai suoi utenti di valutare le opzioni del processo di produzione e i vari compromessi lungo l'intera durata del processo di progettazione. Lo strumento fornisce un'analisi potente e flessibile di costi, tempi di produzione e molti altri fattori di base associati alla produzione e all'assemblaggio del prodotto.

Il programma fornisce in output informazioni preziose per negoziare efficacemente con i fornitori.

3.2.2.2. MicroEstimating

MicroEstimating offre una soluzione di pianificazione dei processi e determinazione dei costi assistita dal calcolatore. [33]

MicroEstimating utilizza diversi strumenti per determinare tempi e costi di produzione:

- Emulatore di machine utensili;
- Riconoscimento automatico delle features;
- Lavorazioni basate sulla conoscenza.

MicroEstimating importa direttamente la distinta materiali (BOM) da SolidWorks, fornendo uno strumento potente ma semplice per la stima degli assiemi, indipendentemente dalla complessità o dal numero di elementi.

3.2.2.3. DFM Concurrent Costing e DFA Product Simplification

Il software DFM Concurrent Costing e DFA Product Simplification di Boothroyd Dewhurst [32] consente agli utenti di generare stime accurate dei costi di parti e assemblati in fase di progettazione. Il software DFM Concurrent Costing permette agli utenti di capire quali sono i principali fattori di costo associati alla produzione del prodotto e stabilisce un punto di riferimento per ciò che il prodotto "dovrebbe costare". Inoltre guida gli utenti attraverso una valutazione di processi e materiali alternativi. I costi si aggiornano automaticamente man mano che gli utilizzatori determinano le tolleranze, le finiture superficiali e altri dettagli delle parti. Man mano che gli utenti scelgono processi di formatura efficaci e considerano come modificare le caratteristiche delle parti per ridurre i costi, il prodotto viene ottimizzato.

3.2.2.4. Costimator

Costimator è un software di stima dei costi sviluppata da Thomas Charkiewicz nel 1982 [34].

Costimator impiega due metodi chiave per la stima del costo del prodotto:

- Modelli di costo parametrici: vengono sviluppati a partire da dati di costo storici. Tempi e costi vengono generati attraverso le tecniche di regressione.
- Stima basata sulle caratteristiche: offre agli utenti con poca o nessuna esperienza di produzione la possibilità di stimare i tempi e i costi in base all'identificazione e alla selezione delle caratteristiche delle parti (ad esempio, fori, fessure, pieghe, ritagli) piuttosto che ai processi di produzione.

In questo caso il programma legge, estrae ed importa automaticamente i dati dei pezzi dai modelli CAD 3D in Costimator.

3.2.2.5. aPriori Product Cost Management

Il Product Cost Management di aPriori [30] offre funzionalità che determinano istantaneamente il costo di una parte o di un prodotto da un modello CAD.

aPriori, infatti, calcola i costi in tempo reale utilizzando le informazioni relative al tipo di materiale, al volume di produzione, al processo di produzione e al luogo di produzione. aPriori può essere eseguito contemporaneamente all'applicazione CAD o come applicazione autonoma.

Determina inoltre il metodo di produzione con il costo più basso per la parte o l'assieme in esecuzione e fornisce tale feedback al progettista in tempo reale.

3.2.3. Strumenti di costo basati sul web

Uno dei siti Web più popolari e utili che supporta le stime dei costi è CustomPartNet [31], una risorsa online per la stima dei costi di produzione. Consente agli utenti di eseguire calcoli rapidi che facilitano la progettazione del prodotto e il processo di costificazione dei manufatti. Con CustomPartNet, gli utenti possono creare rapidamente una nuova stima dei costi o trovare una parte simile nel database delle parti pubbliche, da utilizzare come riferimento. Le stime possono essere salvate e condivise. CustomPartNet ha anche contenuti educativi per aiutare sia gli studenti che gli ingegneri praticanti che sono nuovi nel settore manifatturiero. Le panoramiche dei processi e le linee guida di progettazione consentono agli utenti di esplorare il funzionamento di un processo e di apprendere come progettare le parti in modo più conveniente. Il sito contiene quattro processi di produzione:

- stampaggio a iniezione;
- pressofusione;
- colata in sabbia;
- lavorazione per asportazione di truciolo.

Un altro sito web degno di nota è Get-A-Quote.net [28]. I prezzi derivano da stime da parte di appaltatori che acquistano in volumi moderati e che poi pagheranno ai fornitori che realizzano il pezzo.

3.3. STUDI SU MODELLI DI COSTO PARAMETRICI

La definizione di un approccio per la realizzazione di modelli di costo parametrici per componenti meccanici può essere ottenuta attraverso percorsi, o modalità, che possono essere anche molto differenti tra loro.

Nel caso in esame si è scelto di arrivare ad un modello di costo sfruttando diverse tecniche innovative di machine learning.

Ad oggi l'utilizzo di tali tecniche avviene in diversi contesti, anche di natura molto differente, compreso la stima del costo.

In questo capitolo si vogliono quindi riportare, in maniera concisa, i risultati di alcuni studi caratterizzati dall'uso delle tecniche di machine learning per ottenere stime di variabili dipendenti come il costo, tempi, variabili di processo e altre ancora.

Tsung-Han & S. Deng (2012) [14] svolge uno studio molto interessante nel quale si verifica la fattibilità e l'accuratezza dei metodi di machine learning applicati a due casi di stima dei costi relativi al prodotto, proposti in precedenti studi.

Il primo caso prevede la stima del costo del processo di piegatura per un tubo in acciaio mentre il secondo caso una stima del costo del materiale di un tubo in acciaio al carbonio.

Nel dettaglio il primo caso, come verrà poi fatto in questo caso studio, prevede la raccolta di set di dati suddivisi in campioni di training e di test. Nella fase di training vengono implementate tecniche di reti neurali e tecniche di regressione per stabilire un modello di stima dei costi; successivamente nella fase di test, si osserva come i campioni di prova sono stati immessi nei modelli di stima integrati nella fase di training.

Infine, viene calcolato l'errore tra i valori effettivi e di stima, per valutare le prestazioni dei modelli, dove le metriche utilizzate sono state il MAPE, MSE e R^2 .

La stima dei costi del processo di piegatura dei tubi in acciaio parte da un set di dati di 36 campioni, dove definendo i parametri principali del processo stesso, viene poi suddiviso casualmente in 31 campioni per il training test e 5 per il testing set.

Vengono quindi implementate le reti neurali con un algoritmo di back-propagation (BPN) e la support vector machines (SVM) come tecniche di machine learning, mentre per la regressione vengono utilizzate la tecnica FREG e SREG (rispettivamente first order regression per problemi di predizione semplici e second order regression per problemi di predizione complessi).

I risultati di questo primo caso sono molto interessanti in quanto si evidenzia come le tecniche di machine learning possono ottenere prestazioni di stima più accurate e che il coefficiente di determinazione di questi modelli è superiore al 90%, con un MAPE di 6.16% per la SVM e 8.68% per BPN.

I modelli di regressione invece, hanno dimostrato un minore precisione delle prestazioni in quanto si è capito che per ottenere migliori performance da questo tipo di tecnica si ha bisogno di un campione di training maggiore (risultato a cui anche questo caso studio giunge), dove in questo caso si ottiene un MAPE di 21.49% per SREG e 25.17% per FREG.

Il secondo caso preso in esame invece si pone l'obiettivo di definire un modello di stima dei costi dei materiali per i tubi in acciaio al carbonio. Anche in questo secondo caso la modalità segue perfettamente quella del primo, dove i dati di set sono complessivamente 16 e divisi casualmente in 10 per il training e 6 per la fase di test.

Questo secondo studio conferma i risultati del primo, dove i risultati ottenuti con tecniche di machine learning sono molto più accurate rispetto a quelle ottenute attraverso la regressione; in particolare si ottengono valori di MAPE di 13.23 % per SVM, 11.18 % per BPN e 25.03 % per SREG, 85.87 per FREG.

Da questo studio di fatto si può concludere come le tecniche di machine learning, nella definizione di un modello per la stima dei costi, risultano essere quelle più adatte.

I dati di partenza da cui sono stati ottenuti i risultati della stima di costo dei tubi in acciaio al carbonio, nello studio di Tsung-Han & S. Deng (2012), è di Garza e Rouhana (1995) [15]. Analizzando questo studio, che prevede sempre il confronto tra reti neurali, regressioni lineari e regressioni non lineari si conclude che l'utilizzo di tecniche di machine learning permette di ottenere risultati più soddisfacenti in termini di performance, tuttavia è doveroso considerare che i risultati raggiunti da questo studio potrebbero non essere considerati del tutto corretti, in quanto le regressioni sono state costruite utilizzando l'intero set di 16 osservazioni mentre la rete neurale è stata costruita utilizzando, come descritto in precedenza, 10 osservazioni per il training e 6 per il testing. È anche vero però che i risultati riportati erano errori quadratici medi sull'intero set di addestramento e set.

Osservazioni simili possono essere fatte sullo studio di Brass, Gerrard e Peel (1994) [16] il quale confronta in modo specifico la rete neurale con i modelli di regressione per la stima dei costi di un recipiente a pressione. I risultati ottenuti affermano un miglioramento di un 50% delle performance del modello quando si utilizza una rete neurale rispetto ad un modello di regressione, tuttavia i risultati sono quasi certamente distorti poiché non è stato utilizzato alcun campione di convalida separato, rispetto ai campioni utilizzati per l'allenamento. Un altro studio interessante è quello svolto da Tayefeh Hashemi et al (2020) [17], il quale si prefigge di applicare una metodologia di analisi della stima dei costi basata sul costo diretto ed indiretto dei progetti di costruzione, sulla base di studi e ricerche che propongono stime di costo con tecniche di machine learning.

Nel dettaglio, lo studio va ad esaminare 92 studi differenti con l'obiettivo di definire un modello di costo nelle fasi preliminari l'avviamento di un determinato progetto di costruzione (edifici, ferrovie, strade, etc). Dall'analisi di tutti questi studi è emerso che le tecniche che hanno dato una maggiore soddisfazione sono quelle innovative di machine learning; particolare interesse lo ha suscitato le artificial neural network.

Infatti, si è notato come la ANN risulta essere molto sensibile ai dati di input; poiché questa tecnica di apprendimento automatico è basata sui dati, si nota come essa tende a funzionare in modo più accurato se esiste una grande quantità di dati e un set di dati omogeneo per estrarre le relazioni tra i dati disponibili. D'altra parte, il numero di neuroni di input (noti come fattori di costo) ha un effetto diretto sul malfunzionamento del sistema. Di conseguenza, quando il numero di fattori di costo di input aumenta, la complessità del sistema aumenterà e in caso di stima dei costi di costruzione, ha mostrato che l'accuratezza della stima tende a diminuire.

Caputo et al. (2008) [7] confrontano tra loro la regressione lineare e le reti neurali, valutando le performance nella stima dei costi del processo di produzione di recipienti a pressione di grandi dimensioni.

Questo studio è stato abbastanza significativo a causa dei relativi costi coinvolti e dell'ampia variabilità delle configurazioni e delle dimensioni dei recipienti. I metodi sviluppati sono stati testati su un periodo di sei mesi nelle strutture di uno dei principali produttori mondiali con risultati molto incoraggianti: il metodo di funzione parametrica ha permesso di ottenere un errore di stima medio del 12,5%, con valori estremi all'interno dell'intervallo di variabilità del $\pm 33\%$; la rete neurale si è rilevata ancora più performante e ha permesso di ridurre ulteriormente l'errore medio a meno del 9% con un intervallo di variabilità compreso tra il 33% e il 22%.

Inoltre, il modello ANN è stato testato anche su due recipienti completamente nuovi, di diversa configurazione rispetto a quelle incluse nella banca dati, mostrando errori di stima dei costi entro il 5% dei costi effettivi.

Cavalieri et al. (2004) si concentrano nel settore dell'industria automobilistica [4]. Anche in questo caso, l'obiettivo è stato quello di illustrare i risultati di due diversi approcci, rispettivamente tecniche parametriche lineari e rete neurale artificiale, per la stima dei costi di produzione unitari di un nuovo tipo di dischi freno prodotti da un'azienda manifatturiera italiana. I modelli parametrici lineari e ANN sono stati testati e convalidati confrontando i risultati forniti da questi modelli con i costi effettivi dei venti componenti più rilevanti (dischi grezzi) acquistati o fabbricati dall'azienda.

Le analisi statistiche hanno mostrato la superiorità delle reti neurali rispetto al modello parametrico lineare: l'errore di stima assume il valore massimo di circa il 15% per il modello parametrico, mentre per la rete neurale raggiunge la soglia del 10%, e solo in sei casi è superiore al 5%. I risultati sembrano confermare performance migliori per le reti neurali in questo campo di applicazione, ma non una netta superiorità rispetto all'approccio parametrico più "tradizionale". Il punto più rilevante riguarda la logica intrinseca dei due approcci: mentre l'uso di un modello parametrico richiede la specifica dell'espressione analitica della relazione che collega input e output, ciò non è necessario con una rete neurale. Pertanto, l'ANN è caratterizzata dalla possibilità di determinare autonomamente la forma più appropriata del rapporto, mentre un punto critico, soprattutto nel contesto applicativo specifico, è rappresentato dalla ridotta possibilità di interpretare i dati di output (che è fondamentale per l'"ottimizzazione" delle soluzioni progettuali durante il nuovo processo di sviluppo del prodotto).

Nel campo civile, Puckett si è interessato a stime parametriche dei costi per i progetti di ristrutturazione degli edifici. Sono stati presi come riferimento dati sui costi di cinquanta lavori del sistema dell'Università dell'Alaska da sette sedi del campus [8].

In questo caso il modello è stato sviluppato seguendo un approccio parametrico lineare, senza l'utilizzo e il confronto con tecniche di machine learning.

Il lavoro dimostra l'utilità di questa tecnica di stima dei costi parametrici in quanto permette di facilitare la pianificazione e determinare la fattibilità di progetto.

La rete neurale è stata utilizzata in un secondo studio in cui è stato sviluppato un modello parametrico di stima dei costi per i progetti autostradali. I dati di 18 progetti autostradali costruiti a Terranova negli ultimi cinque anni sono stati utilizzati per formare la ANN. La struttura di un semplice ANN è stata prima simulata utilizzando un programma per fogli di calcolo per fornire una rappresentazione trasparente e semplificata di questa tecnica. Sono stati poi utilizzati diversi approcci per determinare i pesi ottimali del training e quindi la struttura finale della ANN.

Hegazy et al. (1998) [18] hanno dimostrato la praticità dell'utilizzo di programmi per fogli di calcolo nello sviluppo di adeguati modelli ANN da utilizzare nel campo delle costruzioni.

Čeh insieme ad altri autori [19] ha come obiettivo analizzare le prestazioni predittive della tecnica di apprendimento random forest rispetto ai modelli basati sulla regressione multipla per la previsione dei prezzi degli appartamenti. Un set di dati che include 7407 registrazioni di transazioni di appartamenti che fanno riferimento alle vendite immobiliari dal 2008 al 2013 nella città di Lubiana, la capitale della Slovenia, è stato utilizzato per testare e confrontare le prestazioni predittive di entrambi i modelli.

Nella procedura di modellazione sono state prese in considerazione le variabili esplicative disponibili come, ad esempio, le caratteristiche strutturali e di età degli appartamenti, le informazioni ambientali e di vicinato, ecc. Tutte le misure di performance hanno rivelato risultati significativamente migliori per le previsioni ottenute con il metodo forestale casuale (random forest), che conferma la prospettiva di questa tecnica di apprendimento sulla previsione del prezzo dell'appartamento.

In alcuni casi, potrebbe essere utile stimare variabili dipendenti diverse dai costi. I metodi che sono stati spiegati in questo capitolo, non si limitano a fare previsioni soltanto dei costi ma possono essere utilizzati anche per altri fini come, ad esempio, la stima del tempo di lavorazione di un processo meccanico.

È chiaro quindi che i metodi di stima sono utilizzabili in diversi ambiti e con diversi fini.

Inoltre, non è possibile individuare a priori la migliore tecnica da utilizzare tra regressione, tecniche basate su alberi decisionali o tecniche basate su reti neurali ma queste devono essere valutate caso per caso.

In questo lavoro di tesi si andranno a sviluppare modelli di preventivazione del costo per le famiglie tecnologiche di fresatura e tornitura. Il modello trovato sarà quello a maggiore qualità tra tutte le varie tecniche elencate poco sopra.

Un ulteriore passo sarà quello di utilizzare questo modello per provare a costificare alcune nuove componenti non utilizzate nella fase di addestramento dell'algoritmo.

4. METODO

In questo capitolo si illustra il metodo operativo utilizzato ai fini della tesi.

Ricordiamo che lo scopo dell'elaborato è duplice:

1. Determinazione dei parametri che inducono maggiormente in "errore" i software LeanCOST e LeanDESIGNER.
2. Determinazione di un approccio parametrico alla costificazione basato su algoritmi di Machine Learning.

4.1. INDIVIDUAZIONE DEI PARAMETRI A MAGGIOR SENSIBILITÀ DI ERRORE

Nel primo obiettivo si scrive tra virgolette il termine "errore" poiché non si può parlare di una vera e propria imprecisione da parte del programma. Il software, infatti, salvo errori nel riconoscimento di features, va a sviluppare un piano di processo a cui ad ogni operazione associa un costo di setup e un costo di lavorazione. Il risultato è un'analisi puntuale del processo produttivo del componente in esame. A tal proposito bisogna precisare che in genere i vari fornitori utilizzano un procedimento di preventivazione dei costi differente tra loro: nel mercato reale entrano in gioco molte variabili che il software non può considerare non essendo univoche e ben definite. A dimostrazione di ciò basta pensare che se si richiede il preventivo di un determinato "codice X" ad un numero "n" di fornitori, avremo "n" differenti proposte di prezzo.

Si possono considerare come esempio di fattori fuorvianti per il software:

- Metodo di stima del prezzo di vendita: se ci rivolgiamo a fornitori come piccole aziende o artigiani è molto raro che questi costifichino i pezzi per via analitica. È piuttosto diffusa invece una preventivazione "ad esperienza" del capo officina. In questo caso il costo può risultare soggetto a "fattori emozionali" per cui ad esempio una forma più strana potrebbe

essere preventivata ad un prezzo maggiore rispetto ad una forma più tradizionale, anche se il processo in sé risulta identico.

- Tipologia di macchinari disponibili: può capitare che un'officina non disponga di tutte le macchine ottimali per la realizzazione di un certo codice. In quel caso è possibile che il piano di processo "reale" abbia un indice di efficienza ridotto rispetto quello "ipotizzato" dal programma. La minore efficacia del processo produttivo si traduce in una lievitazione del prezzo.
- Assorbimento o meno di alcuni costi: può succedere che un fornitore si faccia carico o meno del costo di investimento per la realizzazione di un pezzo. Per costo di investimento si intende il costo che l'officina deve pagare al proprio operatore per la scrittura del programma CNC utile a realizzare il componente.
- Rischio che è disposto a correre il fornitore: una volta preventivato un costo e comunicato all'ufficio acquisti, il prezzo può sì, subire variazioni, ma entro certi limiti. Per questo motivo, per non andare incontro a una perdita economica, si può avere una costificazione maggiorata da parte del fornitore.

Questi sono solo alcuni degli aspetti che un software non può considerare. Ce ne sarebbero tanti altri ma elencare tutti i possibili fattori di discostamento non è negli scopi del lavoro di tesi.

Nell'elaborato si sono andati ad analizzare circa 140 pezzi forniti da ACMA, attraverso i software LeanCOST e LeanDESIGNER.

Per ciascun software l'analisi è stata fatta con database diversi:

- 1) LeanCOST
 - a) Database Hyperlean V9
 - b) Database Hyperlean V11

- c) Database ACMA
- 2) LeanDESIGNER
 - a) Inserendo solo e soltanto il modello .step (quindi senza considerare filettature, rugosità e tolleranze)
 - b) Inserendo il modello CAD con caratteristiche di rugosità, filettature e tolleranze.

Costruito il database con i risultati di queste analisi, si sono aggiunti i costi reali sostenuti dall'azienda per ciascun codice.

A questo punto si sono sviluppati degli indici che potessero essere significativi per indicare la precisione delle costificazioni dei vari programmi (e con i vari database).

Questa analisi fornisce un valore "globale" sulle prestazioni dei SW, pertanto, ci si è posti l'obiettivo di indagare se ci fossero delle particolari caratteristiche che inducessero, più di altre, i programmi all'errore. Si sono quindi stabiliti degli aspetti geometrici e non di cui tener conto nel database. Implementate queste nuove informazioni, si è sfruttato l'utilizzo del plug-in Excel XLSTAT e attraverso la funzione "matrice delle correlazioni", si è risaliti a quei parametri che rendono più sensibili all'errore LeanCOST e LeanDESIGNER. Una volta determinati "quali" sono le caratteristiche da studiare più approfonditamente, ci si è concentrati a capire il "come" influenzassero l'errore graficando l'andamento dell'errore percentuale in funzione di essi. Per fare ciò si sono sviluppati dei grafici Excel dalla cui lettura è stato possibile estrapolare qualche informazione più dettagliata.

La descrizione degli indici per l'analisi globale, dell'utilizzo della matrice delle correlazioni e dell'interpretazione dei grafici, viene rimandata nel capitolo 6 (Caso Studio).

4.1.1. Analisi dei costi "Should Cost"

L'analisi di Should Cost è un processo che mira a determinare il costo, sulla base di fattori oggettivi quali il costo dei materiali, i costi di produzione, gli overheads ed il guadagno del costruttore [7].

Questa analisi è utile sia quando si acquista una parte da un fornitore sia quando si produce internamente. Infatti, nel primo caso permette di ottenere un vantaggio in fase di contrattazione in quanto si conosce più o meno bene il costo effettivo sostenuto dal fornitore per la produzione della parte. Nel secondo caso, l'analisi permette al team di ingegneri di individuare quali sono le operazioni più gravose a livello di costo e, di conseguenza, concentrare le successive operazioni di re-design, sempre in linea con la fattibilità pratica della modifica e nel rispetto delle tolleranze e performance richieste, esattamente su quelle che risultano essere più impattanti.

L'analisi di Should Cost risulta quindi essere essenziale per la Supply Chain e per il team di ingegneria, specialmente durante la fase di creazione di un nuovo prodotto, ottimizzando il costo dello stesso.

Il livello di dettaglio che si può raggiungere in una analisi di Should Cost dipende dalle capacità e conoscenze di colui che costruisce il modello di costo; una elevata accuratezza richiede la conoscenza del processo produttivo per produrre la parte o le operazioni di assemblaggio, dalle prime operazioni di reperimento delle materie prime fino a requisiti di qualità che devono essere soddisfatti dal prodotto finito.

Generalmente, quest'operazione è svolta dal progettista, in quanto conosce in maniera approfondita il design della parte. Tuttavia, potrebbe non avere conoscenza alcuna di tutta quella serie di operazioni che avvengono in outsourcing, si parla ad esempio di saldature, forgiature, microfusioni etc.

Proprio in quest'area la modellazione Should Cost può aiutare ad ottenere una stima completa dei costi del prodotto.

Gli step di una analisi Should Cost sono riportati in Figura 4.1, [2].

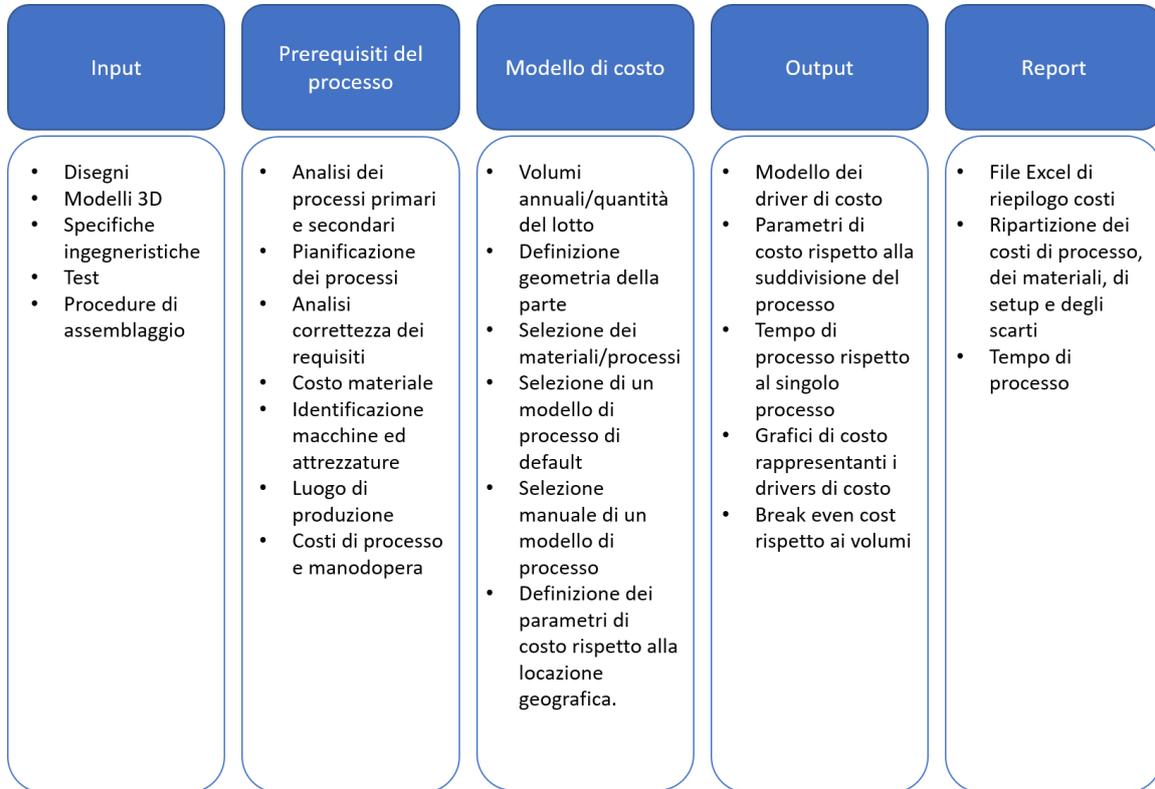


Figura 4.1 - Fasi di un'analisi Should Cost

Per effettuare l'analisi, si considera un prodotto come costituito da una serie di caratteristiche, ognuna delle quali ha una funzione particolare da eseguire.

Ogni prodotto è caratterizzato da:

- Caratteristiche fisiche: i materiali da cui può essere ricavato, il processo di fabbricazione, la precisione, la complessità, le dimensioni e la geometria.
- Caratteristiche economiche: è essenziale conoscere i volumi annuali e le dimensioni del lotto che, a loro volta, aiutano a prendere decisioni di "make or buy".

Il modello di costo si costruisce basandosi sulle caratteristiche del prodotto.

Il progettista può fare un trade-off basato sul comportamento dei costi delle singole funzionalità e ottimizzare il design in base alla criticità della funzionalità.

Oggi sono disponibili software di Should Cost utilizzati per analizzare il comportamento dei costi di una parte o di un assieme. La libreria del software è contrassegnata dalle variabili relative alle caratteristiche del prodotto e del processo ed è costruita sulla base di tecnologie, processi, macchinari, costi orari di manodopera in base a geografie distinte, velocità di configurazione, costi generali e tempi.

La compilazione del modello matematico in un algoritmo impostato comporterà l'osservazione delle variabili di processo che contribuiscono in modo significativo a determinare il costo del prodotto.

Questi software consentono inoltre all'utente di personalizzare materiali e parametri di processo esclusivi non inclusi nella libreria.

Esempi di software per l'analisi di Should Cost sono LeanCOST® e LeanDESIGNER® attraverso i quali è possibile studiare il comportamento dei costi di varie iterazioni di progettazione. Se gli utenti desiderano scegliere un materiale o un processo alternativo, il modello di costo può essere duplicato ed espresso in base alla loro area di interesse per la variazione di progettazione. I grafici statistici consentono all'utente di confrontare e analizzare il comportamento dei costi per le varie iterazioni in termini di processo, materiale, costi generali e volumi di produzione. Riassumendo, le potenzialità di un'analisi Should Cost sono le seguenti:

- Fornisce al team del sourcing una visione migliore su quali siano i margini di profitto dei fornitori
- Permette di identificare i principali driver di costo del prodotto aiutando così nella decisione di "make or buy"
- Permette di effettuare un confronto dei prezzi della parte, o dell'assemblaggio, se esternalizzati a diverse posizioni geografiche;

- Per qualsiasi approvvigionamento strategico, l'analisi dovrebbe essere uno strumento ideale per negoziare proficuamente;
- Consente di analizzare possibili alternative di materiali o processo produttivo per studiarne il costo e fare i giusti trade off;

4.1.2. Software Utilizzato

4.1.2.1. LeanCOST®

Il software LeanCOST® è stato sviluppato da Hyperlean srl.

Lo scopo di questo software è quello di semplificare tutte quelle operazioni da eseguire per determinare il costo di realizzazione di un componente meccanico o di un assieme. Pertanto, una volta che un prototipo CAD 3D è stato sviluppato, il progettista è in grado di definire una stima del costo su cui poi andrà a lavorare il tecnologo. Quest'ultimo ha la possibilità di effettuare modifiche sui parametri tecnologici relativi alle lavorazioni facendo variare le voci di costo (costo materia prima, di set-up, costi accessori e costi delle lavorazioni). L'output di quest'ultima parte è un costo effettivo che verrà poi modificato (o no) dall'ufficio relazioni esterne.

La struttura del programma può essere sommariamente riassunta in quattro moduli [8]:

1. Modulo di interfaccia CAD: analizza il modello CAD e le relative informazioni non geometriche per identificare le caratteristiche costruttive. Il software esegue un'analisi topologica delle entità geometriche (facce, loop, e bordi), dimensioni, finiture, tolleranze e proprietà fisiche (massa e densità);
2. Modulo di allocazione processo: converte l'insieme delle caratteristiche costruttive individuate, in una serie di operazioni per determinare il processo di fabbricazione;

3. Motore di calcolo: calcola automaticamente i tempi di produzione utilizzando le funzioni di calcolo relative ai processi individuati e le traduce in costi;
4. Modulo generazione di report: gestisce i dati calcolati e permette all'utente di utilizzarli in base alle proprie esigenze.

Un punto di forza di questo software è la capacità di integrarsi con i più noti sistemi CAD 3D commerciali e di ricavare autonomamente le caratteristiche geometriche del prodotto e definire i processi di lavorazione necessari.

LeanCOST® è a tutti gli effetti una piattaforma aziendale per la gestione dei costi di produzione e lo scambio di informazioni tra le diverse figure coinvolte attraverso interfacce utente dedicate. Infatti, ogni figura analizza il costo con un diverso livello di dettaglio e dispone degli strumenti necessari a supportare il proprio lavoro in modo efficiente e completo.

Nonostante le numerose potenzialità, questo software non sostituisce il lavoro eseguito dal tecnologo o dal progettista: esso è uno strumento utile per semplificare e velocizzare il lavoro e quindi per ridurre i costi senza però perdere in termini di qualità e precisione. La conoscenza e l'esperienza del tecnologo e del progettista rimangono sempre delle caratteristiche assolutamente indispensabili per la buona riuscita dell'operazione di costificazione.

La logica di funzionamento è la seguente: a partire dal modello CAD 3D e/o da specifiche progettuali, il software analizza la geometria ed estrae i parametri geometrici significativi per ottenere una definizione completa del prodotto in modo automatico. Il motore di calcolo associa quindi la descrizione geometrica ad una o più tecnologie di produzione grazie a regole che racchiudono processi e procedure aziendali (knowledge), estrae i parametri tecnologici e definisce la tecnologia grazie alle informazioni che caratterizzano il contesto specifico (macchine e materie prime).

Il sistema al termine dell'analisi determina per ogni componente il tempo di produzione e il costo, suddiviso in cinque voci di costo che sono: costo della materia prima, costo dell'investimento, di set-up, degli accessori e delle operazioni.

Nonostante la procedura sia automatica, il software lascia ampio spazio alla personalizzazione di qualsiasi particolare della produzione da parte dell'utente. Ciò è reso possibile dall'alto livello di dettaglio per ciascuna voce di costo; è infatti possibile durante la fase di industrializzazione approfondire nel dettaglio tutti i contributi di costo e si può facilmente controllare e, se desiderato, modificare qualsiasi parametro geometrico e tecnologico di qualunque operazione, come ad esempio la velocità di avanzamento dell'utensile di un macchinario o qualsivoglia altra variabile di processo.

4.1.2.2. LeanDESIGNER®

Il software LeanDESIGNER® è un altro software sviluppato da Hyperlean srl. Lo scopo di questo software, al contrario del precedente non è quello di andare a stimare il prezzo finale di realizzazione, bensì nasce per affiancare come strumento che affianca il design engineer nella realizzazione corretta del manufatto guidandolo attraverso le regole del DFMA. In questa tesi, tuttavia, non è stata presa in considerazione questa funzione bensì ci si è accontentati di ottenere da questo strumento una stima del costo del pezzo esaminato. Una volta importato il CAD 3D e inserite le informazioni necessarie per avviare l'analisi quali:

- Materiale
- Logica
- Trattamento termico
- Trattamento superficiale
- Quantità da produrre

Si riceve in output un range di costo del modello più o meno ampio. Al contrario di LeanCOST, il piano di processo non sarà né visibile, né modificabile ad eccezione della tipologia e dimensione della materia prima. L'unica cosa che il tecnologo può fare è eliminare dalla stima del costo alcune features o fasi come, ad esempio, il controllo delle tolleranze.

La struttura del programma può essere sommariamente riassunta in quattro moduli:

1. Modulo di interfaccia CAD: analizza il modello CAD e le relative informazioni non geometriche per identificare le caratteristiche costruttive e gli errori di design secondo le linee guida del DFMA. Il software esegue un'analisi topologica delle entità geometriche (facce, loop, e bordi), dimensioni, finiture, tolleranze e proprietà fisiche (massa e densità);
2. Modulo di allocazione processo e individuazione degli errori: converte l'insieme delle caratteristiche costruttive individuate, in una serie di operazioni e messaggi di errore;
3. Motore di calcolo: calcola automaticamente i costi di produzione delle varie feature utilizzando le funzioni di calcolo relative ai processi individuati;
4. Modulo generazione di report: gestisce i dati calcolati e permette all'utente di utilizzarli in base alle proprie esigenze.

Anche questo software ha come punto di forza la capacità di integrarsi con i più noti sistemi CAD 3D commerciali e lavorare con essi fianco a fianco ricavandone automaticamente le caratteristiche geometriche del prodotto.

La logica di funzionamento è la seguente: a partire dal modello CAD 3D e/o da specifiche progettuali, il software analizza la geometria ed estrae i parametri geometrici significativi per ottenere una definizione completa del prodotto in modo automatico. Il motore di calcolo associa quindi la descrizione geometrica ad uno o più errori di progettazione seguendo le linee guida del DFMA. Il software assegna anche ad ogni feature un range di costo che sommate vanno a generare in output il range del costo finale.

4.2. ANALISI MODELLO PARAMETRICO

Per quanto riguarda il secondo obiettivo, si vuole sviluppare concretamente un modello di costo parametrico da fornire al disegnatore per effettuare una costificazione rapida durante la fase di design del pezzo. Sappiamo dalla letteratura, che per la valutazione di un progetto è fondamentale saperne le potenzialità economiche il prima possibile.

Per lo sviluppo di un modello parametrico si va a riprendere il database sviluppato per la prima analisi e lo si va a rielaborare in modo da ottenere un database tale da essere processato con gli strumenti idonei. Lo strumento in questione è un tool sviluppato internamente in università attraverso visual studio di Microsoft nella piattaforma .NET. Esso è in grado di ricevere in input il database da analizzare (scritto in una certa maniera) per ottenere in output:

- 1 file .zip contenente le informazioni del modello parametrico trovato nell'analisi. Queste informazioni serviranno per implementare nel tool sviluppato anche un'appendice in grado di processare nuovi pezzi non appartenenti al database di partenza.
- 3 file Excel:
 - 1 contenente il database iniziale
 - 1 contenente il database di training

- 1 contenente il database di testing: questo file rappresenta il primo utilizzo del modello di costo. Il tool sviluppa una colonna chiamata “costo predetto dal modello” che sarà confrontata successivamente con la colonna del “costo stimato da LeanCOST” per calcolare gli indici di qualità del modello. In particolare, si sceglie di calcolare i valori di MAPE e R^2 .

4.2.1. Definizione Database per analisi parametrica

Precedentemente si è detto come il database abbia un ruolo importante sul risultato dei modelli di costo.

Esso è costituito da dei parametri indipendenti, ovvero parametri geometrici e non geometrici dei componenti presi per lo studio e dal parametro dipendente, ovvero il parametro che si vuole predire, che in questo studio è il costo.

I parametri indipendenti rappresentano quei parametri che verranno forniti come input al tool per ottenere una stima di costo, pertanto essi devono essere rappresentativi delle principali grandezze che ne possono definire il costo preliminarmente; un esempio potrebbero essere le grandezze geometriche rilevanti, come le dimensioni di ingombro, che possono esseri estratti dalle tavole tecniche dei componenti presi per lo studio oppure grandezze non geometriche come il materiale o il lotto di produzione.

Il parametro dipendente, ovvero il costo, può essere ottenuto andando a riprendere, per i componenti utilizzati per lo studio, i costi effettivamente sostenuti per la loro realizzazione (Actual Cost); oppure nel caso in cui questo non risulta essere possibile, è necessario effettuare una should cost analysis con il supporto di software come LeanCOST®.

Definire la tecnica più adatta allo studio o comunque ottenere un modello di costo più efficiente è strettamente collegato alle caratteristiche principali del database.

Come verrà poi dettagliatamente descritto nel capitolo del caso studio, le caratteristiche principali del database sono principalmente due:

- Numero di parametri indipendenti;
- Numero di records.

Il primo rappresenta, appunto, il numero di parametri indipendenti che si implementa all'interno del database mentre il secondo rappresenta il numero di componenti presi come campionatura per allenare il modello.

Per valutare la loro influenza è possibile creare più database a partire dal database principale.

Il database principale è possibile definirlo come quel database ottenuto andando a considerare tutti i parametri indipendenti che possono essere definiti per quel componente ed un numero di records pari al numero di componenti che si hanno a disposizione per lo studio, i quali forniscono i valori dei parametri indipendenti.

Da questo database principale è possibile ottenere più database che tengano conto di una diversa combinazione di queste due caratteristiche principali.

In particolare, per valutare l'influenza del numero dei parametri indipendenti sono state definite due varianti: la prima che conserva un numero di parametri pari a quelli definiti inizialmente per il database principale; la seconda, invece, va a considerare un numero di parametri ridotto, con l'ottica di andare ad eliminare quei parametri che non hanno una forte correlazione con il parametro dipendente e che risultano essere ridondanti tra gli stessi parametri indipendenti, cioè andare ad utilizzare solo quelli strettamente necessari.

Di fatto con questa suddivisione si valuta come il numero di parametri che si utilizza va ad influenzare le performance del modello di costo.

Per valutare la caratteristica numero di records, anche in questo caso si definiscono due varianti: la prima dove il numero di records è esattamente pari a quello del database principale ovvero pari al numero di componenti presi per lo studio; la seconda, invece, risulta essere determinata dalla diversa combinazione di alcuni parametri indipendenti per lo stesso componente. Ad esempio, è possibile per lo stesso componente valutare diverse opzioni di lotto, ciò comporta ad un valore diverso del parametro dipendente, che eseguito per più componenti, ne accresce la campionatura dando maggiori possibilità di allenare il modello con valori differenti.

Si è quindi in presenza di due caratteristiche suddivise in due varianti ciascuna, le quali combinandole permettono di valutare come le stesse possono influire sul risultato.

Nella Figura 4.2 si evidenziano le possibili combinazioni.

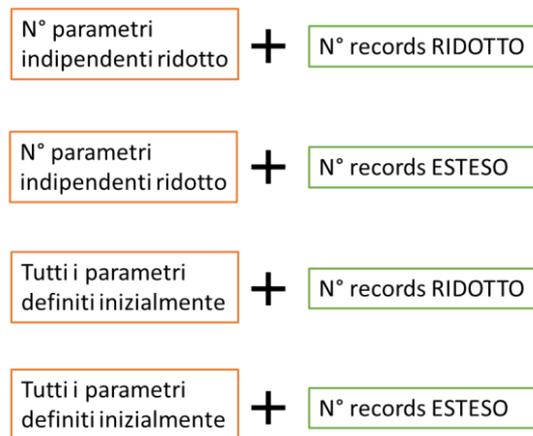


Figura 4.2 - Combinazioni delle varianti delle caratteristiche principali del database

Ottenuti quindi i diversi database rappresentanti le differenti combinazioni delle caratteristiche principali, segue l'operazione di split data, infatti, non tutti i dati vengono utilizzati per generare un modello ma ci sarà sempre una restante parte dedicata a una fase di test. Sono proprio i test-set che permettono di capire come risponde un modello e quindi validare i risultati.

In particolare, una buona percentuale di suddivisione varia dal 60% all'80% di dati di training e da 40% a 20% di dati di testing.

4.2.2. Software Utilizzato

Il tool utilizzato in questa analisi, come già detto è stato sviluppato internamente in università attraverso visual studio di Microsoft nella piattaforma .NET.

Tale strumento, una volta che riceve in pasto il database definito nella fase precedente, effettua in sequenza le operazioni per dare in output i file utili all'analisi.

Le operazioni fatte in sequenza dal tool sono quelle riportate in Figura 4.3:

- Analisi Pearson: si eliminano le caratteristiche che influenzano in maniera marginale la variabile dipendente
- Split data: si divide in maniera randomizzata il database di partenza in "training set" e "test set" nelle proporzioni rispettivamente del 80% e 20%.
- Analisi di questi database attraverso varie metodologie: Le metodologie implementate nel tool sono:
 - Regressione Lineare;
 - Rete Neurale;
 - Random Forest;
 - Gradient Boosting;
- Confronto tra le analisi e scelta del modello migliore: ogni metodologia avrà una certa qualità del risultato misurata dai parametri R^2 e MAPE. Essendo quattro le metodologie implementate nel tool, il programma esegue l'analisi con ciascuna di esse ottenendo dunque quattro coppie di valori R^2 e MAPE. A questo punto il tool confronta questi valori scegliendo il modello a più alta qualità.

- Export del modello: scelto il modello ottimale, il programma crea un file .zip che contiene la formula/l'albero decisionale/la rete neurale ottenuta dall'analisi. Questo dipende da quale metodo viene scelto come ottimale dal tool. Si utilizzerà questo file per implementare nel tool uno strumento in grado di processare nuovi pezzi non appartenenti al database di partenza.
- Export analisi: oltre al file zip del punto precedente il tool va a creare tre file Excel:
 - Database iniziale completo.
 - Database usato per il training.
 - Database usato per il testing.

Nel foglio Excel con il database di testing il programma usa il modello trovato per sviluppare una colonna di "costo predetto" in base agli input del database di test.

Andando a confrontare il costo predetto dal modello con quello del database si può trovare la qualità del modello trovato andando a vedere i parametri di R^2 e MAPE.

Riassumendo posso dire dunque di avere per il tool:

- INPUT:
 - Database iniziale
- OUTPUTS:
 - 1 file .zip che servirà per implementare nel tool la funzione di costificazione da usare su codici non appartenenti al DB iniziale.
 - 3 file Excel da cui in quello di testing set è possibile leggere la qualità del modello.

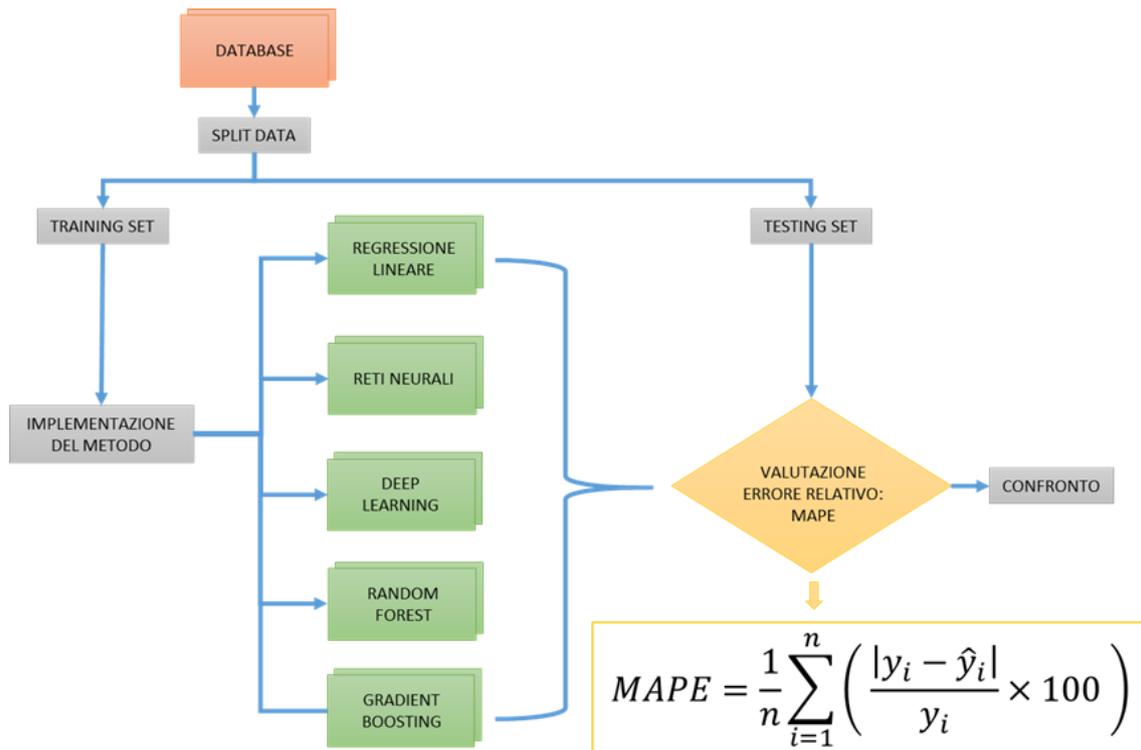


Figura 4.3 - Metodologia per lo sviluppo modelli di costo parametrici

4.2.3. Analisi dei risultati

Dopo aver utilizzato il software per determinare il modello parametrico che più riesce ad avvicinarsi ad una costificazione corretta se ne valuta la qualità. Per fare ciò si tengono in considerazione alcuni indici come il MAPE (Mean Absolute Percentage Error) e R^2 .

5. CASO STUDIO: PARTE 1

5.1. PREPARAZIONE DEL DATABASE

Per il lavoro di tesi sono stati forniti dall'azienda circa 130 componenti dei quali, escludendo i pezzi di cui non si avevano abbastanza informazioni, o modificati durante il processo di progettazione se ne ottengono 118 utilizzabili per l'analisi. Di questi 118 componenti 51 riguardano la famiglia tecnologica dei fresati mentre i restanti 67 appartengono alla famiglia tecnologica dei torniti.

La costruzione del database consiste nell'andare a costificare tutte le componenti fornite dall'azienda attraverso i programmi LeanCOST e LeanDESIGNER.

Per ciascun programma inoltre bisogna fare delle precisazioni.

Per quanto riguarda LeanCOST, si sono dovute utilizzare due differenti versioni per un totale di tre database diversi:

- LeanCOST ver.9 – Database Hyperlean
- LeanCOST ver.9 – Database ACMA
- LeanCOST ver.11 – Database Hyperlean

Si è fatta tale distinzione poiché era utile indagare sulla precisione della previsione in ciascun caso e determinare quale fosse il database ottimale.

Per quanto riguarda invece LeanDESIGNER si è scelto di usare tale strumento in due modi distinti:

- Inserendo tutte le varie tolleranze, filettature e rugosità;
- Dando in pasto al programma il CAD grezzo (senza le caratteristiche del punto sopra).

Si precisa che per ciascun codice si è valutato con i due software sia il prezzo unitario che quello a quantità reale d'acquisto. In particolare, per LeanCOST si è potuto definire il prezzo per i diversi range come in Tabella 5.1

Q.tà	1	2	3	4	5-7	8-10	11-15	16-20	21-31	32-61
€/pezzo	aa	bb	cc	dd	ee	ff	gg	hh	ii	mm

Tabella 5.1 - Tabella dei costi unitari per diversi range di quantità

Per LeanDESIGNER invece, si è eseguita la preventivazione sulla quantità unitaria e sulla quantità acquistata dall'azienda.

Q.tà	1	n
€/pezzo	kk	ww

Tabella 5.2 - Tabella dei costi unitari a quantità unitaria (1) e quantità d'acquisto reale (n)

Durante la fase di costificazione veniva annotato il tempo necessario all'analisi.

Si sono considerati:

- Tempo di analisi della tavola e trascrizione delle informazioni necessarie sul CAD;
- Tempo di analisi tramite LeanDESIGNER (unico per i due modi di utilizzo);
- Tempo di analisi tramite LeanCOST (unico per i tre diversi database).

Una volta ottenuti questi dati, in base alle quantità del lotto dei diversi codici si sono filtrati i costi individuati dai diversi software e dai diversi database in base alla quantità d'acquisto. Tali dati vengono poi affiancati dalla colonna dei "costi reali" sostenuti dall'azienda per quel tipo di manufatto per andare a fare un confronto.

Si ottiene quindi una tabella del tipo seguente:

Codice	€/pezzo	€/pezzo	€/pezzo	€/pezzo	€/pezzo	€/pezzo
	LC V9	LC ACMA	LC V11	LD base	LD	"costo reale"
xxxxx	yy	zz	ww	rr	dd	tt

Tabella 5.3 - Tabella riassuntiva dei costi stimati

Compilate tutte le righe con queste informazioni e volendo indagare sulla presenza dei parametri che rendono i programmi più sensibili all'errore, si definiscono alcune caratteristiche geometriche (e non).

Si scelgono a tal proposito:

- Famiglia tecnologica principale: fresatura, tornitura;
- Famiglia di materiale: acciai al carbonio, acciai inox e allumini;
- Dimensioni di ingombro del finito [mm]: D_1, D_2, D_3 ;
 - Rapporti tra le dimensioni: $\frac{D_1}{D_2} ; \frac{D_3}{D_1} ; \frac{D_3}{D_2}$
- Tipologia di grezzo: piatto, tondo, lamiera, tubo circolare ecc.;
- Volume di ingombro [mm³];
- Volume del finito [mm³];
 - Rapporto Volume finito/ Volume di ingombro (in %);
- Tolleranza più spinta in classe IT e μm corrispondenti;
- Finitura superficiale più spinta;
- Finitura superficiale standard;
- Dimensione superficie a rugosità bassa [mm²];
- Dimensione superficie a rugosità standard;
 - Rapporto superficie a rugosità bassa/ superficie a rugosità standard (in %);
- % del costo di setup sul prezzo totale;
- Numero di utensili necessari;
- Numero di centri di costo stimati.

In Appendice A viene rappresentato un esempio della tabella risultante.

Completata questa analisi preliminare il database è pronto per essere analizzato dunque si procede con la fase successiva.

5.2. ANALISI EFFICIENZA GLOBALE

Prima di indagare su quali caratteristiche del pezzo inducono più all'errore i programmi utilizzati, si determinano le prestazioni globali dei diversi software/database per identificare quale di essi sia il migliore.

Per ogni tipo di analisi si definiscono i seguenti indici prestazionali [Tabella 5.4]:

- Δ Costo: sta ad indicare la differenza tra costo reale e costo preventivato. Un valore maggiore di zero indica che il programma sovrastima il prezzo, mentre un valore minore di zero che sottostima;
- Δ Costo (%): sta ad indicare l'errore percentuale rispetto al prezzo reale. Un valore maggiore di zero indica che il programma sovrastima il prezzo, mentre un valore minore di zero che sottostima;
- Valore assoluto del Δ Costo (%): sta ad indicare l'errore percentuale "puro" rispetto al prezzo reale.

Δ C LD base	Δ C (%) LD base	Δ C (%) LD base	Δ C LD	Δ C (%) LD	Δ C (%) LD (val.ass)	Δ C LC v9	Δ C (%) LC v9	Δ C (%) LC v9 (val.ass)	Δ C LC ACMA	Δ C (%) LC ACMA	Δ C (%) LC ACMA (val.ass)	Δ C LC v11	Δ C (%) LC v11	Δ C (%) LC v11 (val.ass)
qq	ww	ee	rr	tt	yy	uu	ii	oo	pp	aa	ss	dd	ff	gg

Tabella 5.4 - Parametri prestazionali per analisi globale

Con questi indici si ricavano:

- Totale Δ Costo: indica lo scostamento totale del prezzo preventivato rispetto a quello realmente sostenuto dall'azienda se tutti i pezzi analizzati fossero acquistati in quantità unitaria;

$$\Delta C_{TOT} = \sum \Delta C_i$$

- Media pesata degli errori percentuali: il peso viene dato dalla percentuale sul costo totale del pezzo preso in esame (considerando il prezzo unitario)

$$M_{pesata} = \frac{\sum Err. \%_i * \% \text{ su costo tot a qtà unit.}_i}{\sum \% \text{ su costo tot a qtà unit}}$$

- Media pesata degli errori percentuali in valore assoluto: il peso viene dato come nel caso precedente.

$$M_{pesata(ass)} = \frac{\sum Err. \% (ass)_i * \% \text{ su costo tot a qtà unit.}_i}{\sum \% \text{ su costo tot a qtà unit}}$$

Concettualmente, pesare l'errore, sta a significare che una discrepanza di un ipotetico 10% su un pezzo che costa 100 valga di più sulla stima del budget rispetto allo stesso errore del 10% su un pezzo che costa 10. La differenza che consegue dal considerare l'errore percentuale con segno e in valore assoluto differenzia i due indici in quanto:

- Il primo, poiché considera il segno, potrebbe dare un risultato ingannevole nel caso ci fossero errori grandi ma ben distribuiti in valori minori e maggiori di zero.
- Con il secondo indice si va invece a dare un'informazione sull'errore "puro".

Esempio:

Se si prendessero i valori nella Tabella 5.5:

Val. #1	Val. #2	Val. #3	Val. #4
80%	80%	-80%	-80%

Tabella 5.5 - Valori arbitrari per esempio indici

considerando il loro peso identico gli indici risulterebbero i seguenti:

$$M_{pesata} = 0\%$$

$$M_{pesata(ass)} = 80\%$$

Si può subito intuire che un valore basso di M_{pesato} non per forza indica precisione da parte del software. Al contrario, $M_{pesato(ass)}$ riesce a rappresentare l'efficienza del software in maniera più precisa.

- Δ Costo % totale: questa volta si considera ciascun codice nella quantità effettiva di acquisto e non in quantità unitaria. Può essere interpretato come la differenza in termini monetari tra i costi stimati dal programma e i costi realmente sostenuti dall'azienda.

$$\Delta C(\%)_{TOT} = \frac{C_{TOT SW}}{C_{TOT REALE}} * 100$$

5.3. INDAGINE SUI PARAMETRI

Una volta determinata la precisione a livello globale del software ci si pone l'obiettivo di individuare se esistano delle caratteristiche geometriche (e non) che inducano più o meno in errore i vari software. Per fare ciò ci si avvale dei dati descritti nel Paragrafo 4.1.

Si identificano dunque, attraverso un'analisi Pearson, i valori delle variabili indipendenti a più alta correlazione con la variabile dipendente esaminata. Come variabili indipendenti si considerano le informazioni di tipologia del grezzo, tipologia di materiale, dimensioni di ingombro ecc. mentre come variabile dipendente si utilizza il $\Delta C(\%)_{ass}$, considerato il parametro più rappresentativo dell'errore "puro" del software.

5.3.1. Analisi Pearson

L'analisi Pearson, è un'analisi che si basa sul concetto di correlazione, ovvero sulla relazione esistente tra due o più variabili.

Lo scopo di questo tipo di analisi è quello di andare a diminuire il numero dei parametri indipendenti definiti inizialmente.

Ovviamente, la riduzione di questi parametri deve essere eseguita con una certa logica; tale logica si pone l'obiettivo di andare a valutare l'intensità del legame tra i parametri, attraverso il calcolo di un indice di correlazione.

In questo studio è stato considerato come indice di correlazione la "r" di Pearson, anche detto coefficiente di correlazione.

La caratteristica principale di questo coefficiente è che il suo valore varia nell'intervallo [-1;+1]; tali valori estremi rappresentano relazioni perfette tra le variabili, mentre 0 rappresenta l'assenza di relazione. Questo almeno finché si considerano relazioni di tipo lineare.

Più pragmaticamente, lo studio di correlazione dei parametri mira ad eliminare quei parametri che secondo l'analisi di correlazione risultano essere maggiormente correlati ad altri parametri, considerandoli come ridondanti.

Per poter mettere in pratica questa analisi si è sfruttato il plug-in di Excel XLSTAT, che permette proprio di eseguire questo tipo di analisi.

Il primo step è avviare il plug-in, andare sulla sezione "descrizione dei dati" e selezionare "Matrici di similarità/dissimilarità (correlazione)".

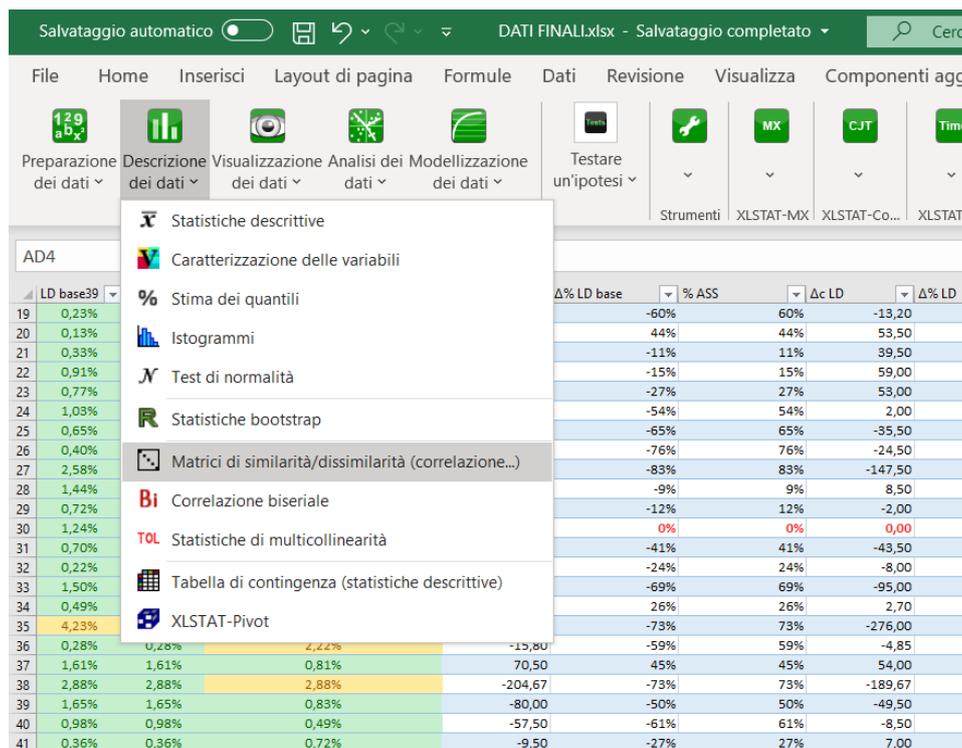


Figura 5.1 - Inizializzazione del comando "Matrice di correlazione"

Avviata la funzione, si seleziona il database da esaminare.

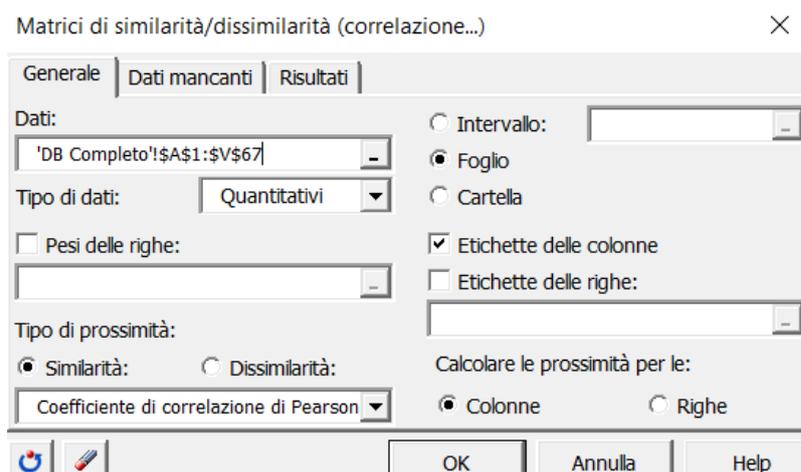


Figura 5.2 - Selezione del database

In output avremo la *“correlation matrix”* ovvero una matrice nel quale si indica con un numero compreso tra -1 e +1 la relazione tra le variabili. Di questa matrice se ne trova una seconda composta dai valori assoluti di quest’ultima, Tabella 5.6:

NOTE	logia matellogia del gr	D1 [mm]	Spess. [mm]	D3 [mm]	o materia prime grezzo [ne fitito [r app. Volumena + spirititura + spissa ngostigostia ed i geometrici Rapp. Sup	D3/D1	costo di set	utensili	Le centri di co	CLASSE	Indipenden																				
NOTE	1 0,062792	0,195777	0,179347	0,257824909	0,443463436	0,333750151	0,374362	0,407341	0,143642	0,374239	0,165655	0,297	0,208028	0,288749	0,013664	0,19075	0,301033	0,523387	0,485066	0,566835	0,121868										
Tipologia r	0,062792	1 0,098295	0,246133	0,397589477	0,012677498	0,250271918	0,19748	0,183448	0,166527	0,170098	0,233377	0,333856	0,179537	0,309268	0,247949	0,007051	0,149383	0,384308	0,098094	0,34144	0,272384										
Tipologia c	0,195777	0,098295	1 0,099782	0,19877109	0,047190658	0,083709966	0,064542	0,013856	0,15902	0,448028	0,06917	0,011367	0,0992108	0,105463	0,083229	0,11446	0,124053	0,157306	0,009593	0,535043	0,083369										
D1 [mm]	0,179347	0,246133	0,099782	1 0,859232366	0,002379783	0,757509714	0,773821	0,685578	0,335653	0,133379	0,238972	0,683509	0,631803	0,354583	0,350047664	0,511985573	0,732317667	0,680594	0,72804	0,090195	0,191758	0,010371	0,572781	0,756081	0,307497	0,023445	0,002256	0,52583	0,015383	0,640734	0,339238
Spess. [mm]	0,257825	0,397589	0,198771	1 0,859232366	0,002379783	0,757509714	0,773821	0,685578	0,335653	0,133379	0,238972	0,683509	0,631803	0,354583	0,350047664	0,511985573	0,732317667	0,680594	0,72804	0,090195	0,191758	0,010371	0,572781	0,756081	0,307497	0,023445	0,002256	0,52583	0,015383	0,640734	0,339238
D3 [mm]	0,443463	0,012677	0,047191	0,00238	1 0,11258156	0,311382449	0,257151162	0,97131	0,83796	0,233996	0,245231	0,163788	0,676202	0,744324	0,255812	0,139089	0,121399	0,732318	0,570346	0,148458	0,631429	0,148703									
Costo mat	0,33375	0,250272	0,08371	0,75751	0,683438026	0,311382449	0,257151162	0,97131	0,83796	0,233996	0,245231	0,163788	0,676202	0,744324	0,255812	0,139089	0,121399	0,732318	0,570346	0,148458	0,631429	0,148703									
Volume gr	0,374239	0,170098	0,448028	0,135379	0,168904637	0,23182302	0,24523938	0,229694	0,18337	0,406207	0,406207	0,703804	0,420902	1 0,420902	0,244867	0,077022	0,313589	0,183471	0,047382	0,191758	0,49416	0,389252	0,380721	0,323759							
Volume fir	0,407341	0,183448	0,013836	0,685578	0,537744054	0,46056142	0,837960142	0,875816	0,069531	1 0,18337	0,154169	0,724222	0,889233	0,280152	0,136967	0,072791	0,72864	0,628041	0,132961	0,663366	0,157834										
Rapp. Vol.	0,143642	0,166527	0,15902	0,335653	0,370978912	0,150018484	0,233996048	0,231527	0,069531	1 0,406207	0,703804	0,420902	1 0,420902	0,244867	0,077022	0,313589	0,183471	0,047382	0,191758	0,49416	0,389252	0,380721	0,323759								
Tolleranza	0,374239	0,170098	0,448028	0,135379	0,168904637	0,23182302	0,24523938	0,229694	0,18337	0,406207	0,406207	0,703804	0,420902	1 0,420902	0,244867	0,077022	0,313589	0,183471	0,047382	0,191758	0,49416	0,389252	0,380721	0,323759							
Finitura + i	0,165655	0,233377	0,06917	0,238972	0,221592342	0,114666845	0,163787582	0,160154	0,154169	0,703804	0,420902	1 0,420902	0,244867	0,077022	0,313589	0,183471	0,047382	0,191758	0,49416	0,389252	0,380721	0,323759									
Sup. a basi	0,297	0,033856	0,011367	0,683509	0,39607754	0,296881622	0,744323829	0,735251	0,8883233	0,004163	0,077022	0,142328	1 0,142328	0,13121	0,670138	0,670138	0,160581	0,182681	0,275161	0,275161	0,100875	0,307544	0,002256	0,52583	0,015383	0,640734	0,339238				
Sup. a ruga	0,208028	0,179537	0,092108	0,631803	0,38296857	0,513349418	0,744323829	0,735251	0,8883233	0,004163	0,077022	0,142328	1 0,142328	0,13121	0,670138	0,670138	0,160581	0,182681	0,275161	0,275161	0,100875	0,307544	0,002256	0,52583	0,015383	0,640734	0,339238				
Prezenza d	0,288749	0,309268	0,105463	0,354583	0,408005418	0,161558825	0,255811969	0,26341	0,280152	0,303122	0,133589	0,292328	0,614871	0,070513	0,572781	0,54831	0,145485	0,150454	0,457553	0,005447	0,129768	0,339238	0,311095	0,403786	0,331802						
Rapp. Sup	0,013664	0,247949	0,133827	0,233884	0,210551603	0,023407072	0,13908895	0,124691	0,136967	0,492571	0,183471	0,614871	0,070513	0,572781	0,54831	0,145485	0,150454	0,457553	0,005447	0,129768	0,339238	0,311095	0,403786	0,331802							
D3/D1	0,19075	0,007051	0,11446	0,42419	0,377034291	0,605846516	0,12139908	0,161384	0,072791	0,501862	0,047382	0,269827	0,070513	0,572781	0,54831	0,145485	0,150454	0,457553	0,005447	0,129768	0,339238	0,311095	0,403786	0,331802							
% costo di	0,301033	0,149383	0,124053	0,64994	0,530047664	0,511985573	0,732317667	0,680594	0,72804	0,090195	0,191758	0,010371	0,572781	0,54831	0,145485	0,150454	0,457553	0,005447	0,129768	0,339238	0,311095	0,403786	0,331802								
n° utensili	0,523387	0,384308	0,157306	0,575748	0,585212086	0,294330803	0,570345608	0,600978	0,628041	0,43363	0,49416	0,399242	0,529725	0,468406	0,655953	0,28196	0,133672	0,52583	0,015383	0,640734	0,339238										
n° centri di	0,485066	0,098094	0,083229	0,009593	0,108465478	0,295585014	0,148457918	0,175792	0,132961	0,143599	0,389252	0,197435	0,089037	0,001085	0,162208	0,046027	0,171262	0,015383	0,323531	0,403786	0,087919										
CLASSE	0,566835	0,31414	0,108975	0,535043	0,516291416	0,534490139	0,631428977	0,621642	0,663366	0,246042	0,380721	0,323964	0,54831	0,542469	0,636523	0,195716	0,070442	0,640734	0,786098	0,403786	0,331802										
V. Indipen	0,121868	0,272384	0,083369	0,173289	0,161603351	0,151450031	0,148703032	0,140542	0,157834	0,020894	0,323759	0,027908	0,145485	0,150454	0,457553	0,005447	0,129768	0,339238	0,311095	0,403786	0,331802										

Tabella 5.6 - Matrice delle correlazioni completa

Una volta arrivati a questo punto si utilizza la formattazione condizionale per evidenziare i valori più importanti dell'analisi in base al valore risultante del *coefficiente di correlazione r*.

In particolare, seguendo le informazioni in letteratura [12] [13], si decide di formattare con:

- $r < 0,3$: valori di bassa correlazione con l'errore percentuale (variabile dipendente) → Giallo;
- $r > 0,8$: valori di elevata correlazione tra le variabili dipendenti → Rosso.

A questo punto il primo passaggio consiste nell'eliminare quei parametri a bassa correlazione con la variabile dipendente, con $r < 0,3$. Dei valori rimanenti si vanno ad escludere quei valori strettamente correlati tra loro, ovvero con $r > 0,8$. Nel caso di due o più variabili reciprocamente correlate si sceglie di eliminare quelle la cui sommatoria degli "r di Pearson" risulta più alta.

Effettuati tutti questi passaggi si ottiene la matrice finale dove rimangono solo e soltanto quei parametri maggiormente legati alla variabile dipendente. Un esempio di matrice ottenuta la possiamo vedere nella Tabella 5.7 che segue:

	Tolleranza + spinti	geometrico	costo di set	n° utensili LC	CLASSE	V. Indipendente
Tolleranza + spinta [μm]	1	0,313589	0,191758	0,49416	0,380721	0,32375874
Presenza di geom. fuori asse	0,313589	1	0,307497	0,655953	0,636523	0,457552705
% costo di setup	0,191758	0,307497	1	0,52583	0,640734	0,339237502
n° utensili LC	0,49416	0,655953	0,52583	1	0,786098	0,311095375
CLASSE	0,380721	0,636523	0,640734	0,786098	1	0,331802431
V. Indipendente	0,323759	0,457553	0,339238	0,311095	0,331802	1

Tabella 5.7 - Matrice delle correlazioni ridotte

Dei parametri indipendenti rimasti si deve in seguito studiare l'influenza sulla variabile dipendente.

Questo esempio esplicativo mostra il procedimento utilizzato per ridurre il numero di parametri indipendenti. Tale procedura deve essere applicata ad ogni database.

Nel caso studio specifico i parametri ottenuti dall'analisi per i differenti database e le differenti famiglie tecnologiche vengono riportati di seguito:

- Tornitura:

LD base	-----	-----	-----	-----	-----
LD	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	n° utensili LC	Classe
LC 11	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	n° utensili LC	Classe
LC 9	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	n° utensili LC	Classe
LC ACMA	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	Classe	

Tabella 5.8 - Parametri ad alta correlazione per la famiglia tecnologica della tornitura

- Fresatura:

Per la fresatura, dopo aver svolto l'analisi Pearson, non si riesce ad evidenziare alcuna correlazione tra la percentuale di errore del software e i parametri identificati per l'analisi.

Questo può essere interpretato come una maggiore casualità dell'errore che riguarda questa famiglia tecnologica.

Un esempio di risultato ottenuto dall'analisi Pearson è il seguente:

Come si può vedere tutti i parametri hanno un “*r di Pearson*” minore di 0,3 rispetto alla variabile dipendente.

5.3.2. Studio del comportamento della variabile in funzione degli indici rimanenti

Attraverso l’analisi Pearson riusciamo a filtrare solo i parametri strettamente correlati all’errore percentuale ma non riusciamo a capire in che modo queste voci possano influenzare la variabile dipendente. Per studiare questo bisogna andare a graficare l’errore percentuale in funzione dei parametri ottenuti nel sottoparagrafo 5.3.1.

Si ricorda che per la fresatura, nessun indice è abbastanza significativo in correlazione con l’errore percentuale della stima, pertanto, si passa all’analisi sui parametri della tornitura.

5.3.2.1. LD base

Per LD base, come già preannunciato nella Tabella 5.8 non ci sono parametri che sensibilizzano più di altri l’errore del software.

Per questo programma si può notare però come l’errore percentuale rispetto ai pezzi di classe D sia sensibilmente inferiore a quello degli altri software.

$$\% \text{ Errore}_{medio_D} = 55\%$$

Si tiene conto di questo risultato per l’analisi dei risultati.

5.3.2.2. LD

- Tolleranza più spinta in μm .

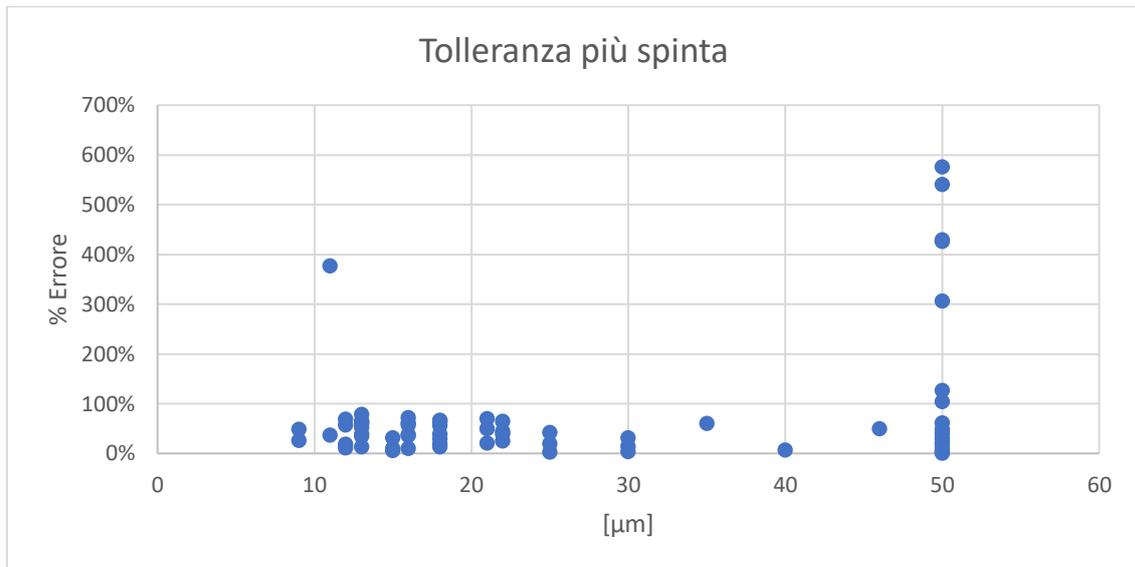


Figura 5.3 - Dati Tolleranza - % Errore - LD

Nel grafico sono rappresentate sull'asse x il valore in micrometri della tolleranza più spinta mentre sull'asse y il valore dell'errore percentuale del software. A tutti i pezzi che non hanno alcuna tolleranza dimensionale si assegna il valore di $50\mu\text{m}$. Dal grafico si nota subito una nube di punti per valori che vanno da $8\mu\text{m}$ a $30\mu\text{m}$. Questa nube di punti si estende per un ampio range di valori sull'asse y ($0\% < \% \text{ Errore} < 80\%$) senza avere però un andamento preciso all'aumentare o al diminuire dell'intervallo di tolleranza. Se si escludono infatti i pochi punti anomali con errore maggiore del 300% si può dire che i dati sono ben distribuiti nel range appena citato.

- Presenza di geometrie fuori asse

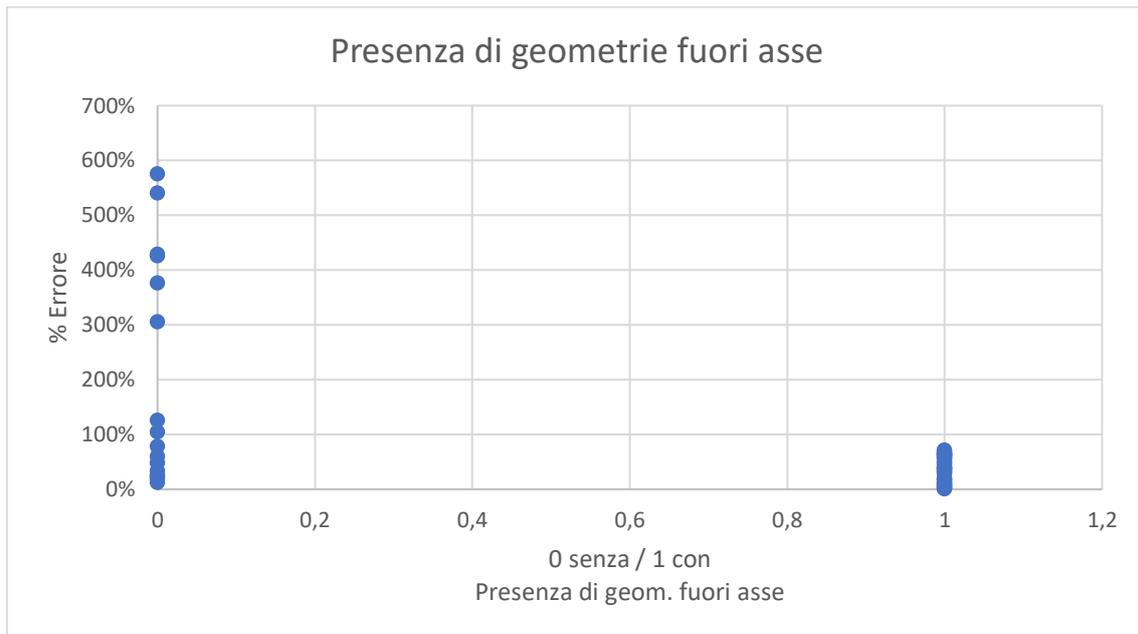


Figura 5.4 - Influenza della presenza di geometrie fuori asse sulla precisione di LD

Nel grafico si indica sull'asse x la presenza (1) o l'assenza (0) di geometrie fuori asse, mentre nell'asse y il valore dell'errore percentuale del software.

Visualizzando i risultati può sembrare che la presenza di geometrie fuori asse, e quindi l'aumentare della difficoltà produttiva del pezzo, possa rendere la previsione del software più precisa. Bisogna però riflettere su questi risultati: in genere se un pezzo non presenta una geometria complessa risulta più semplice da produrre pertanto il prezzo di acquisto risulta essere minore.

Riprendendo il database è possibile vedere che per i pezzi di tipo "0", ovvero quelli che non presentano geometrie fuori asse hanno prezzi molto bassi rispetto ai pezzi di tipo "1". Il costo d'acquisto medio tra i due insiemi è molto diverso come è possibile vedere nella Figura 5.5.

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	7
Codice 7	21
Codice 8	42
Codice 9	2
Codice 10	2
Codice 11	5
Codice 12	25,5
Codice 13	5
Codice 14	12,5
Codice 15	31,5
Codice 16	250
Codice 17	35
Media costo	28,94

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	37
Codice 2	13
Codice 3	15
Codice 4	22
Codice 5	75
Codice 6	100
Codice 7	63
•	•
•	•
•	•
Codice 42	245
Codice 43	185
Codice 44	145
Codice 45	148
Codice 46	97
Codice 47	485
Codice 48	850
Codice 49	710
Media	202,54

(B)

Figura 5.5 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)

I valori medi dell'errore percentuale per le due tipologia sono:

$$\% \text{ Errore}_{medio_0} = 111\%$$

$$\% \text{ Errore}_{medio_1} = 44\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore}_{ass_medio_0} = 111\% * 28,94€ = 32,12€$$

$$\text{Errore}_{ass_medio_1} = 44\% * 202,54€ = 89,12€$$

Si conclude dicendo che all'aumentare della complessità del pezzo, quindi all'insorgere di presenza di geometrie fuori asse, aumenta l'errore assoluto commesso dal software.

- % costo di setup

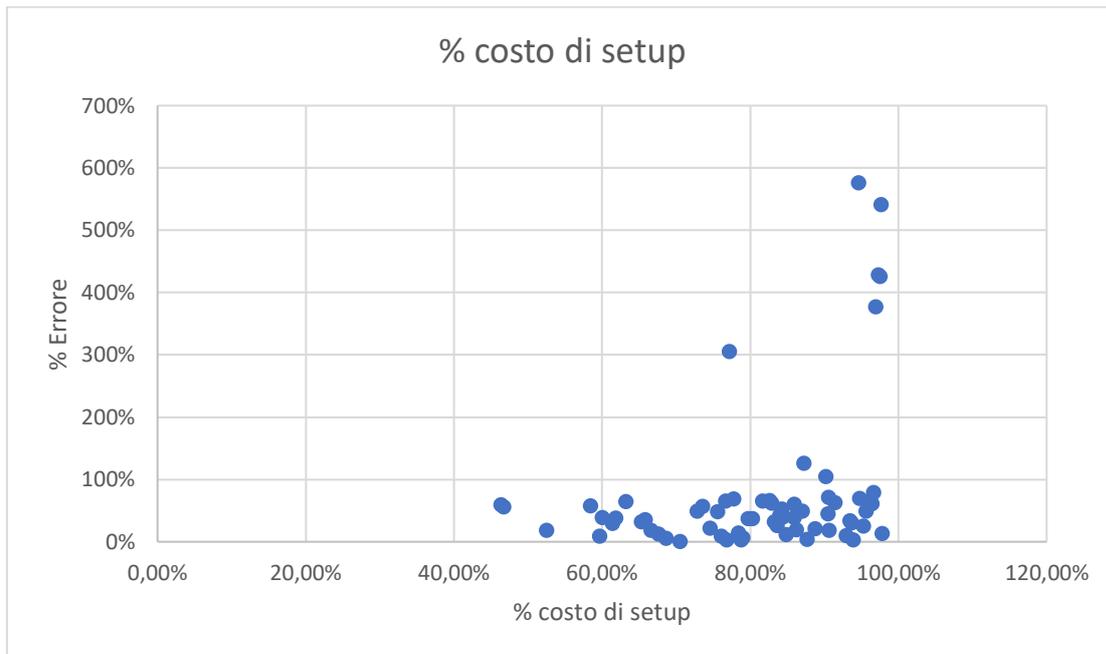


Figura 5.6 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale - LD

Se consideriamo l'andamento dell'errore percentuale in funzione della percentuale del costo di setup sul prezzo totale, abbiamo una nuvola di punti concentrata tra un errore del 2% al 80% a qualunque valore di % costo setup. Tale nuvola però non sembra seguire una certa tendenza.

Si può notare però, che per codici con % del costo di setup maggiore del 94% possono esserci grossi errori nella stima del costo.

I codici che subiscono queste anomalie sono in genere pezzi di semplice realizzazione e dal costo molto basso, fatto per cui non ci si stupisce dall'aver un errore percentuale molto alto.

- N° di utensili stimanti

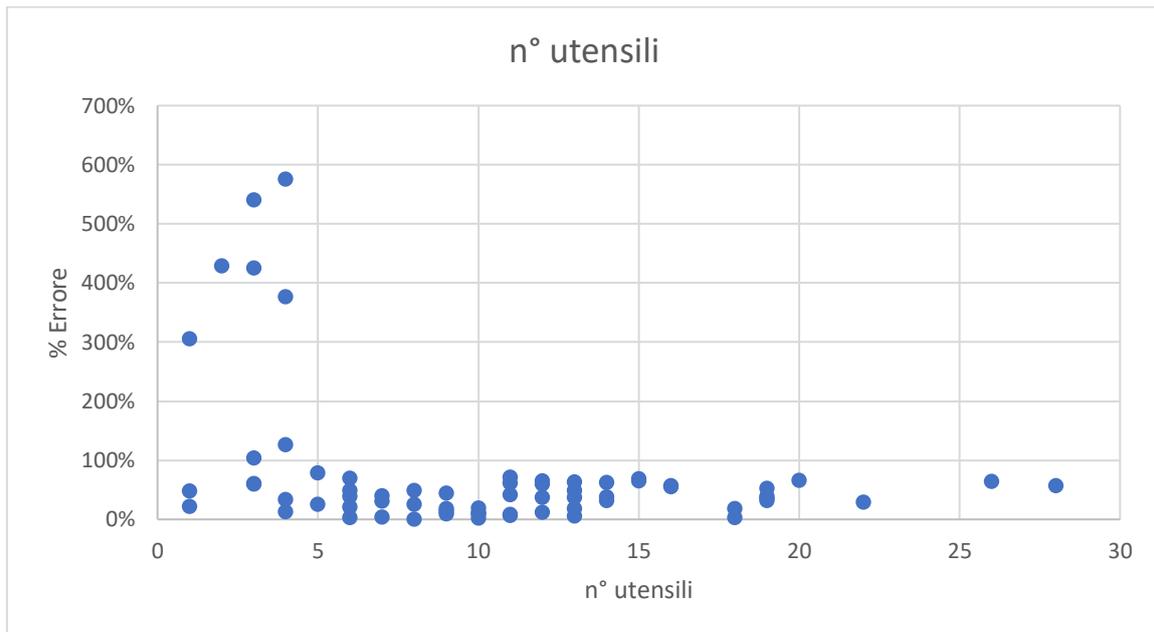


Figura 5.7 - Andamento dell'errore percentuale in funzione del n° di utensili stimati - LD

Andamento dell'errore percentuale in funzione del n° di utensili stimati. Da questo grafico si può vedere come quando il numero di utensili è basso l'errore varia in un range più ampio rispetto al caso in cui il numero di utensili stimato cresce. Questo andamento può essere spiegato dicendo che il n° di utensili è "proporzionale" al prezzo di acquisto del prodotto. Sbagliare di tanto rispetto a un pezzo che necessita di 5 utensili è più facile che sbagliare di tanto rispetto ad un pezzo che necessita di 20 utensili.

Nel seguito si prendono in considerazione due gruppi di codici:

1. Gruppo che necessita da 1 a 10 utensili;
2. Gruppo che necessita da 11 a 28 utensili.

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7
•	•
•	•
•	•
Codice 28	139
Codice 29	150
Codice 30	22,1
Codice 31	140
Codice 32	135,5
Codice 33	85,5
Codice 34	96
Codice 35	165
Media	58,49

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	75
Codice 2	63
Codice 3	39
Codice 4	140
Codice 5	67,5
Codice 6	145
Codice 7	26,85
•	•
•	•
•	•
Codice 24	245
Codice 25	185
Codice 26	145
Codice 27	148
Codice 28	97
Codice 29	485
Codice 30	850
Codice 31	710
Media	269,98

(B)

Figura 5.8 - Costo medio per pezzi con n° utensili da 1 a 10 (A) e da 11 a 28 (B)

$$\% \text{ Errore_medio}_{1 \rightarrow 10} = 79\%$$

$$\% \text{ Errore_medio}_{11 \rightarrow 28} = 35\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore_ass_medio}_{1 \rightarrow 10} = 79\% * 58,49\text{€} = 46,20\text{€}$$

$$\text{Errore_ass_medio}_{11 \rightarrow 28} = 35\% * 269,98\text{€} = 94,49\text{€}$$

Si ottiene una conclusione simile all'analisi fatta precedentemente sulla presenza di geometrie fuori asse: più il pezzo è complesso più l'errore assoluto risulta ampio. Inoltre, se la stima del numero di utensili risulta bassa la probabilità di avere un errore percentuale grande è maggiore.

- Classe di appartenenza

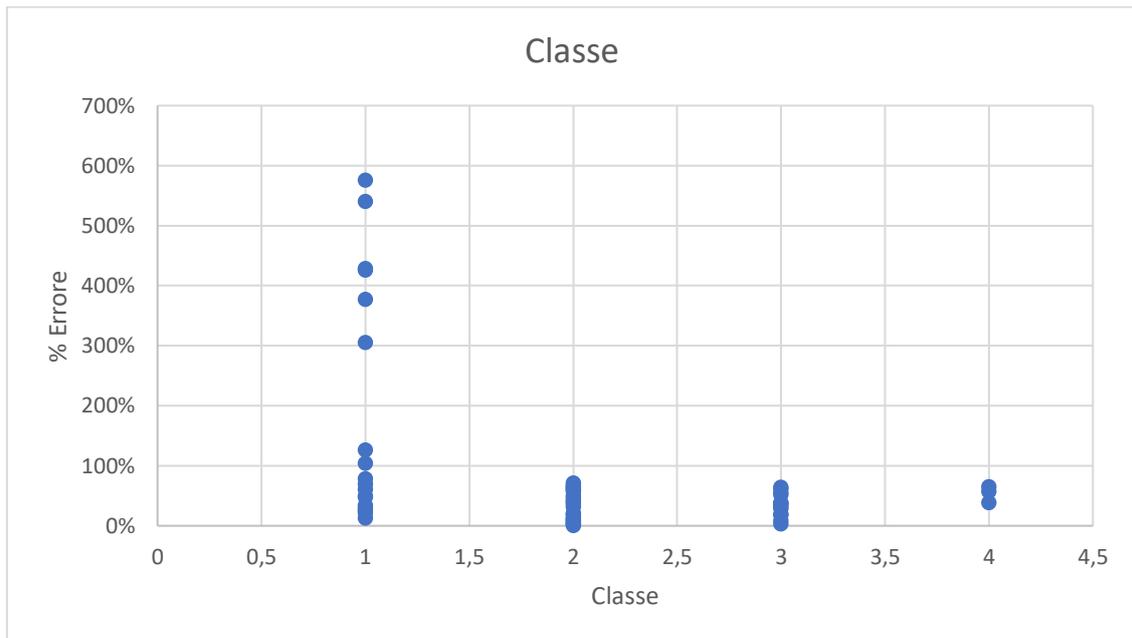


Figura 5.9 - Errore percentuale per le varie classi di produzione - LD

La determinazione della classe di appartenenza è il parametro più soggettivo di quelli analizzati in quanto esso viene determinato attraverso un'ispezione visiva dal responsabile del reparto "cost engineering".

La classificazione non ha particolari canoni per cui inserire un pezzo in una classe piuttosto che in un'altra. Possiamo dire che il responsabile assegna un codice ad una classe in base all'esperienza direzionata dalla complessità di produzione del pezzo.

La classe più semplice è la classe D, seguita dalla C, dalla B e dalla A che risulta la più complessa. Nella Figura 5.9 si indica con il numero 1 la classe D, con il 2 la classe C, con il 3 la classe B e con il 4 la classe A.

Analizzando il costo medio di ciascuna classe e la percentuale di errore medio della stessa è possibile studiarne le conclusioni in maniera più precisa:

$$\% \text{ Errore}_{\text{medio}_D} = 175\%$$

$$\% \text{ Errore}_{\text{medio}_C} = 34\%$$

$$\% \text{ Errore}_{\text{medio}_B} = 34\%$$

$$\% \text{ Errore}_{\text{medio}_A} = 53\%$$

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7
Codice 8	21
Codice 9	42
Codice 10	2
Codice 11	13
Codice 12	2
Codice 13	5
Codice 14	25,5
Codice 15	5
Codice 16	15
Codice 17	22
Codice 18	12,5
Codice 19	31,5
Media	15,47

(D)

Codice	Prezzo d'acquisto
Codice 1	580
Codice 2	580
Codice 3	100
Codice 4	760
Codice 5	620
Codice 6	660
Codice 7	235,5
Codice 8	165
Codice 9	245
Codice 10	185
Codice 11	145
Codice 12	148
Codice 13	97
Media	347,73

(B)

Codice	Prezzo d'acquisto
Codice 1	75
Codice 2	100
Codice 3	63
Codice 4	39
Codice 5	250
Codice 6	140
Codice 7	70
Codice 8	67,5
Codice 9	21
Codice 10	145
Codice 11	26,85
Codice 12	156
Codice 13	279,17
Codice 14	160
Codice 15	95
Codice 16	35
Codice 17	141
Codice 18	98
Codice 19	129
Codice 20	139
Codice 21	150
Codice 22	195
Codice 23	22,1
Codice 24	140
Codice 25	190
Codice 26	140
Codice 27	135,5
Codice 28	53,45
Codice 29	85,5
Codice 30	96
Media	114,569

(C)

Codice	Prezzo d'acquisto
Codice 1	485
Codice 2	850
Codice 3	710
Media	681,67

(A)

Figura 5.10 - Costo medio delle varie Classi

$$Errore_{ass_medio_D} = 117\% * 15,47€ = 18,09€$$

$$Errore_{ass_medio_C} = 34\% * 114,57€ = 38,95€$$

$$Errore_{ass_medio_B} = 34\% * 347,73€ = 118,23€$$

$$Errore_{ass_medio_A} = 53\% * 681,67€ = 361,29€$$

Possiamo concludere che il valore assoluto dell'errore aumenta all'aumentare della complessità del pezzo, in accordo con i dati ottenuti dalle analisi precedenti.

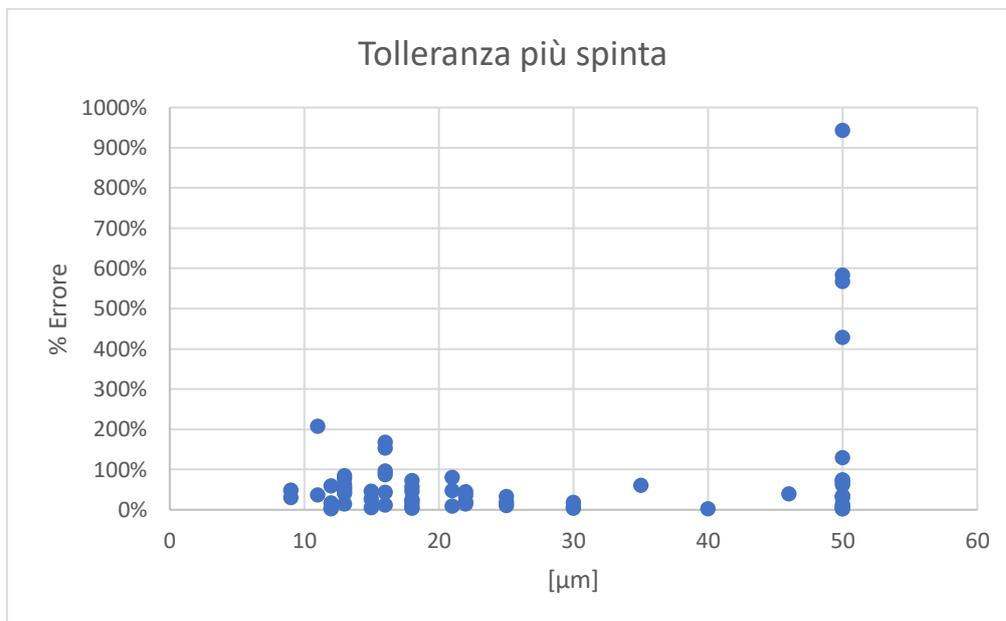
Si può notare come per pezzi di classe A, oltre all'errore assoluto si ha un aumento anche dell'errore percentuale. I dati su quest'ultima classe derivano sull'analisi di soli tre codici, dovrebbero perciò essere integrati con ulteriori analisi per avere un valore statisticamente significativo.

Per il pezzi di classe D la probabilità di avere un errore percentuale alto è maggiore.

Si prosegue sviluppando la stessa analisi per i vari database di LeanCOST.

5.3.2.3. LeanCOST 11

- Tolleranza più spinta in μm .



- Presenza di geometrie fuori asse

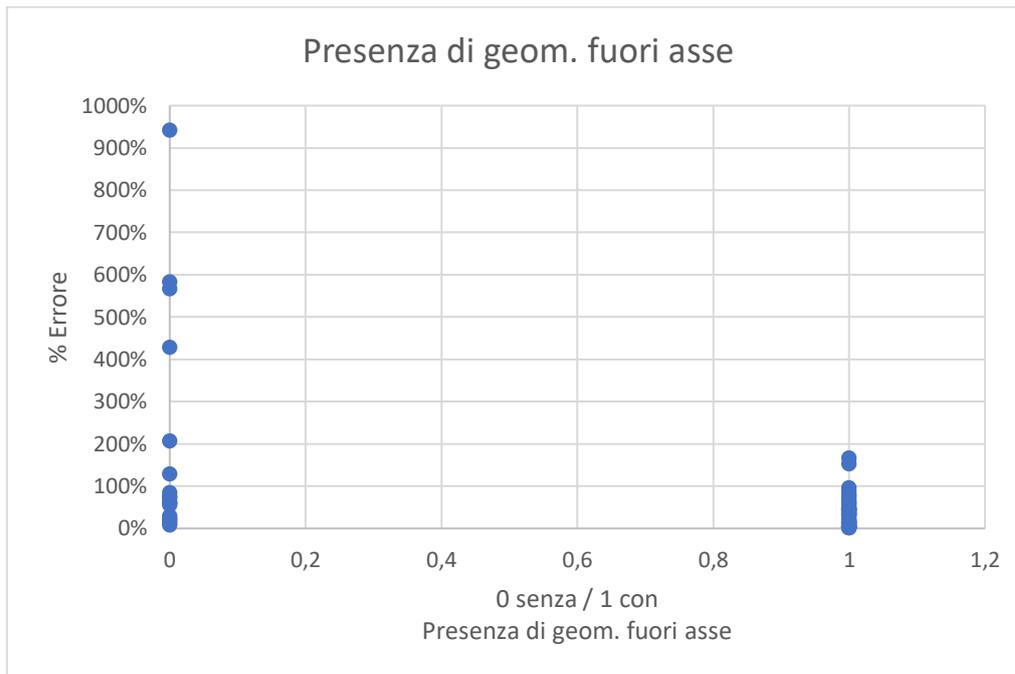


Figura 5.12 - Influenza della presenza di geometrie fuori asse sulla precisione di LC11

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	7
Codice 7	21
Codice 8	42
Codice 9	2
Codice 10	2
Codice 11	5
Codice 12	25,5
Codice 13	5
Codice 14	12,5
Codice 15	31,5
Codice 16	250
Codice 17	35
Media costo	28,94

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	37
Codice 2	13
Codice 3	15
Codice 4	22
Codice 5	75
Codice 6	100
Codice 7	63
•	•
•	•
•	•
•	•
Codice 42	245
Codice 43	185
Codice 44	145
Codice 45	148
Codice 46	97
Codice 47	485
Codice 48	850
Codice 49	710
Media	202,54

(B)

Figura 5.13 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)

I valori medi dell'errore percentuale sono:

$$\% \text{ Errore}_{\text{medio}_0} = 118\%$$

$$\% \text{ Errore}_{\text{medio}_1} = 45\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore}_{\text{ass}_{\text{medio}_0}} = 118\% * 28,94\text{€} = 34,15\text{€}$$

$$\text{Errore}_{\text{ass}_{\text{medio}_1}} = 45\% * 202,54\text{€} = 91,14\text{€}$$

- % costo di setup

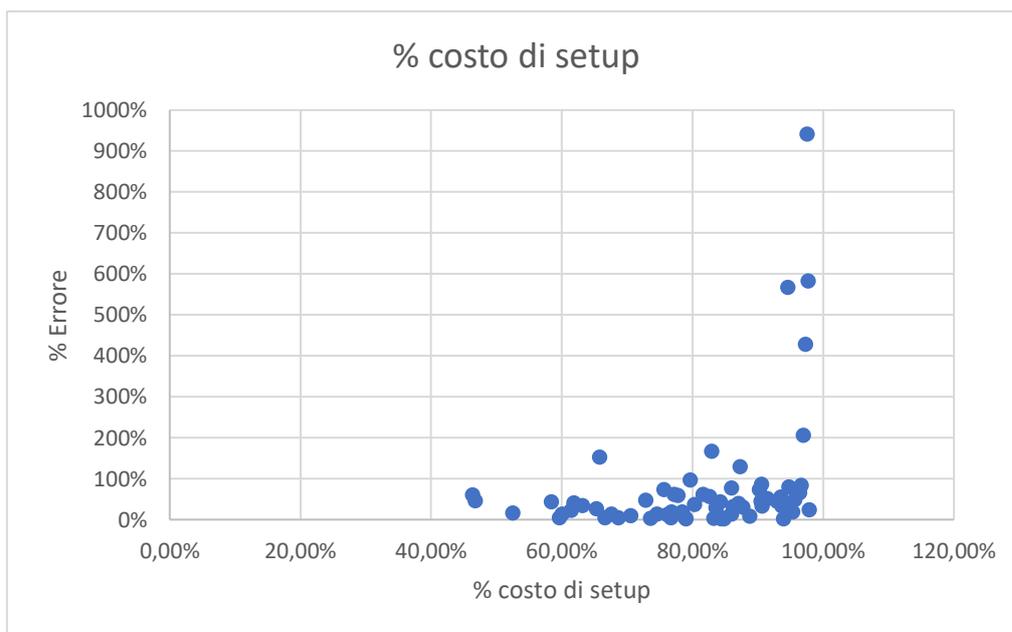


Figura 5.14 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale – LC11

Si possono ripetere le considerazioni fatte per LeanDESIGNER nel capitolo 5.3.2.2.

- N° di utensili stimanti

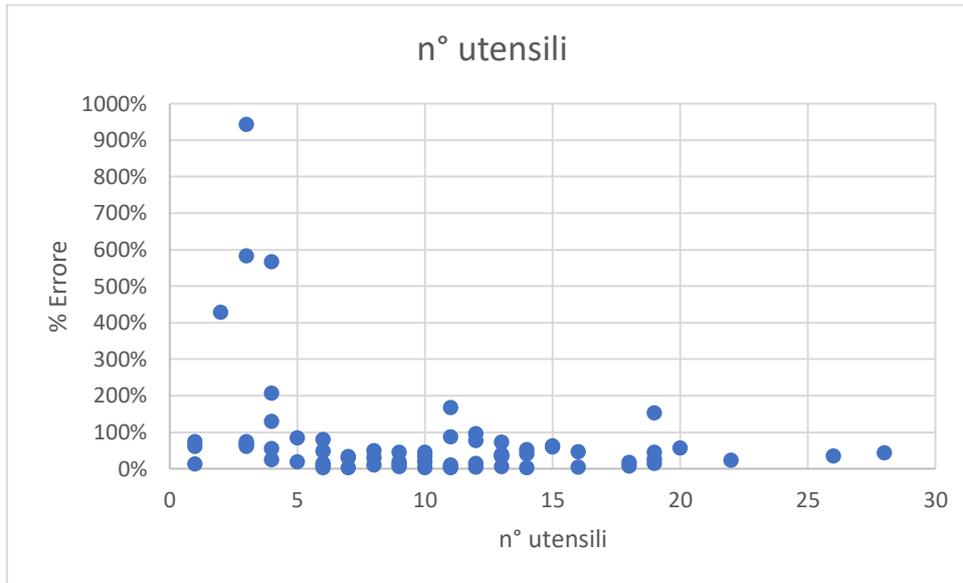


Figura 5.15 - Andamento dell'errore percentuale in funzione del n° di utensili stimati – LC11

Da questo grafico si può vedere come al crescere del numero di utensili il range di errore percentuale decresce.

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7

• •
• •
• •

Codice 28	139
Codice 29	150
Codice 30	22,1
Codice 31	140
Codice 32	135,5
Codice 33	85,5
Codice 34	96
Codice 35	165
Media	58,49

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	75
Codice 2	63
Codice 3	39
Codice 4	140
Codice 5	67,5
Codice 6	145
Codice 7	26,85

• •
• •
• •

Codice 24	245
Codice 25	185
Codice 26	145
Codice 27	148
Codice 28	97
Codice 29	485
Codice 30	850
Codice 31	710
Media	269,98

(B)

Figura 5.16 - Costo medio per pezzi con n° utensili da 1 a 10 (A) e da 11 a 28 (B)

$$\% \text{ Errore}_{\text{medio}_{1 \rightarrow 10}} = 84\%$$

$$\% \text{ Errore}_{\text{medio}_{11 \rightarrow 28}} = 35\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore_ass_medio}_{1 \rightarrow 10} = 84\% * 58,49\text{€} = 49,13\text{€}$$

$$\text{Errore_ass_medio}_{11 \rightarrow 28} = 35\% * 269,98\text{€} = 94,49\text{€}$$

- Classe di appartenenza

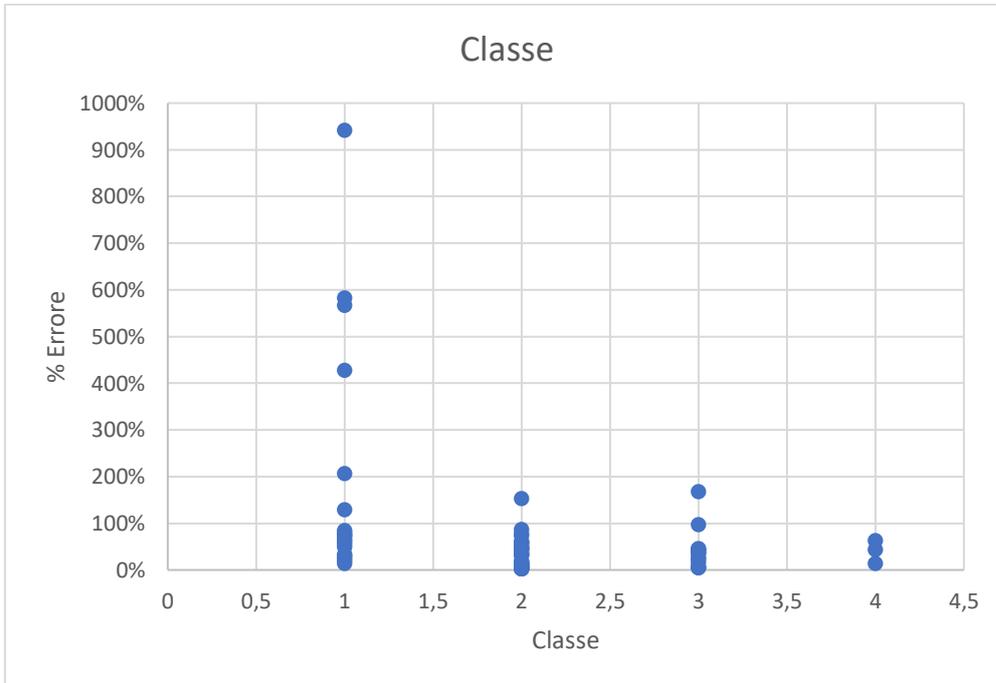


Figura 5.17 - Errore percentuale per le varie classi di produzione – LC11

Gli errori percentuali di ciascuna classe sono:

$$\% \text{ Errore_medio}_D = 184\%$$

$$\% \text{ Errore_medio}_C = 32\%$$

$$\% \text{ Errore_medio}_B = 41\%$$

$$\% \text{ Errore_medio}_A = 39\%$$

Gli errori assoluti medi di ciascuna classe risultano:

$$\text{Errore_ass_medio}_D = 184\% * 15,47\text{€} = 28,46\text{€}$$

$$\text{Errore_ass_medio}_C = 32\% * 114,57\text{€} = 36,66\text{€}$$

$$\text{Errore_ass_medio}_B = 41\% * 347,73\text{€} = 142,57\text{€}$$

$$\text{Errore_ass_medio}_A = 39\% * 681,67\text{€} = 265,85\text{€}$$

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7
Codice 8	21
Codice 9	42
Codice 10	2
Codice 11	13
Codice 12	2
Codice 13	5
Codice 14	25,5
Codice 15	5
Codice 16	15
Codice 17	22
Codice 18	12,5
Codice 19	31,5
Media	15,47

(D)

Codice	Prezzo d'acquisto
Codice 1	580
Codice 2	580
Codice 3	100
Codice 4	760
Codice 5	620
Codice 6	660
Codice 7	235,5
Codice 8	165
Codice 9	245
Codice 10	185
Codice 11	145
Codice 12	148
Codice 13	97
Media	347,73

(B)

Codice	Prezzo d'acquisto
Codice 1	75
Codice 2	100
Codice 3	63
Codice 4	39
Codice 5	250
Codice 6	140
Codice 7	70
Codice 8	67,5
Codice 9	21
Codice 10	145
Codice 11	26,85
Codice 12	156
Codice 13	279,17
Codice 14	160
Codice 15	95
Codice 16	35
Codice 17	141
Codice 18	98
Codice 19	129
Codice 20	139
Codice 21	150
Codice 22	195
Codice 23	22,1
Codice 24	140
Codice 25	190
Codice 26	140
Codice 27	135,5
Codice 28	53,45
Codice 29	85,5
Codice 30	96
Media	114,569

(C)

Codice	Prezzo d'acquisto
Codice 1	485
Codice 2	850
Codice 3	710
Media	681,67

(A)

Figura 5.18 - Costo medio delle varie Classi

Possiamo concludere che per il valore assoluto dell'errore aumenta all'aumentare della complessità del pezzo come detto per LD.

5.3.2.4. LeanCOST 9

- Tolleranza più spinta in μm .

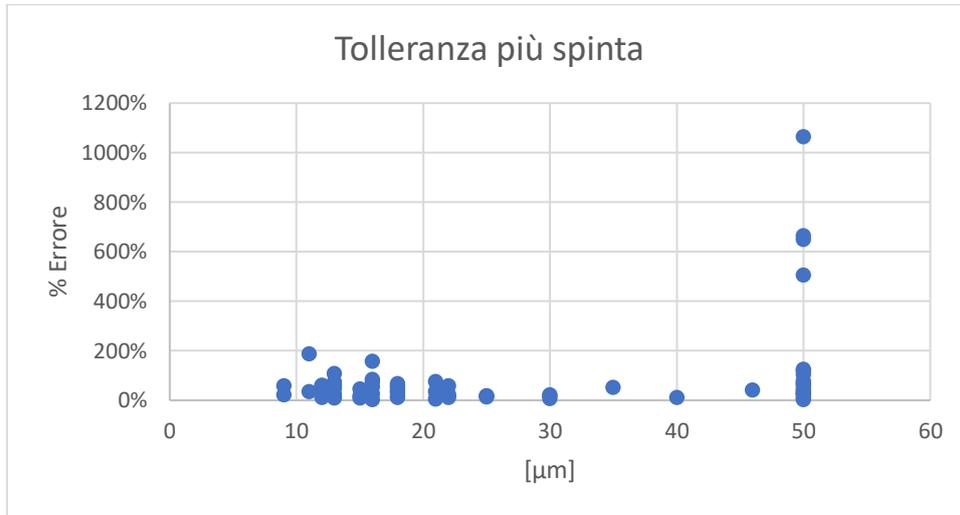


Figura 5.19 - Dati Tolleranza - % Errore - LC9

Si possono ripetere in maniera identica le considerazioni fatte nel capitolo 5.3.2.2.

- Presenza di geometrie fuori asse

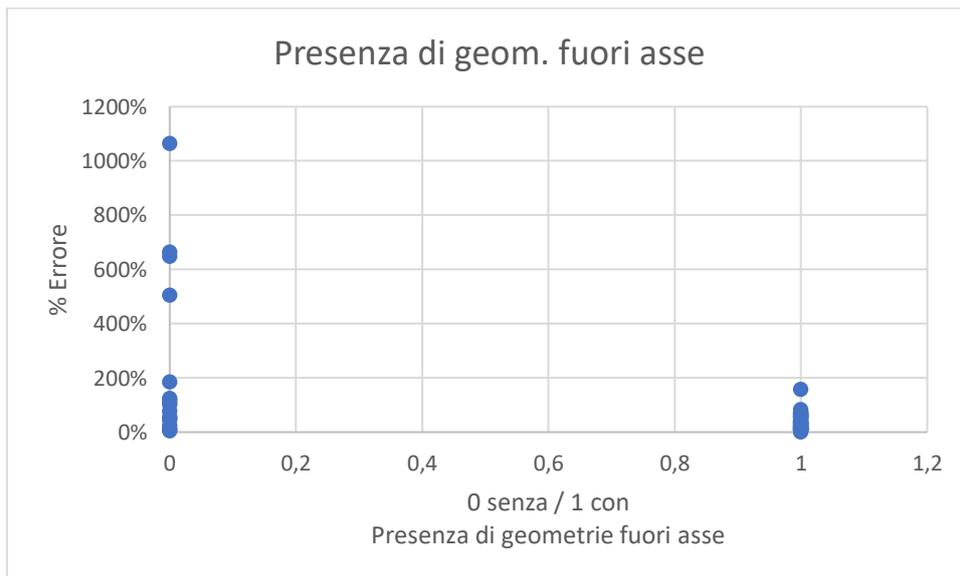


Figura 5.20 - Influenza della presenza di geometrie fuori asse sulla precisione di LC9

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	7
Codice 7	21
Codice 8	42
Codice 9	2
Codice 10	2
Codice 11	5
Codice 12	25,5
Codice 13	5
Codice 14	12,5
Codice 15	31,5
Codice 16	250
Codice 17	35
Media costo	28,94

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	37
Codice 2	13
Codice 3	15
Codice 4	22
Codice 5	75
Codice 6	100
Codice 7	63
•	•
•	•
•	•
Codice 42	245
Codice 43	185
Codice 44	145
Codice 45	148
Codice 46	97
Codice 47	485
Codice 48	850
Codice 49	710
Media	202,54

(B)

Figura 5.21 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)

I valori medi dell'errore percentuale sono:

$$\% \text{ Errore}_{\text{medio}_0} = 129\%$$

$$\% \text{ Errore}_{\text{medio}_1} = 46\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore}_{\text{ass}_{\text{medio}_0}} = 129\% * 28,94€ = 37,33€$$

$$\text{Errore}_{\text{ass}_{\text{medio}_1}} = 46\% * 202,54€ = 93,17€$$

- % costo di setup

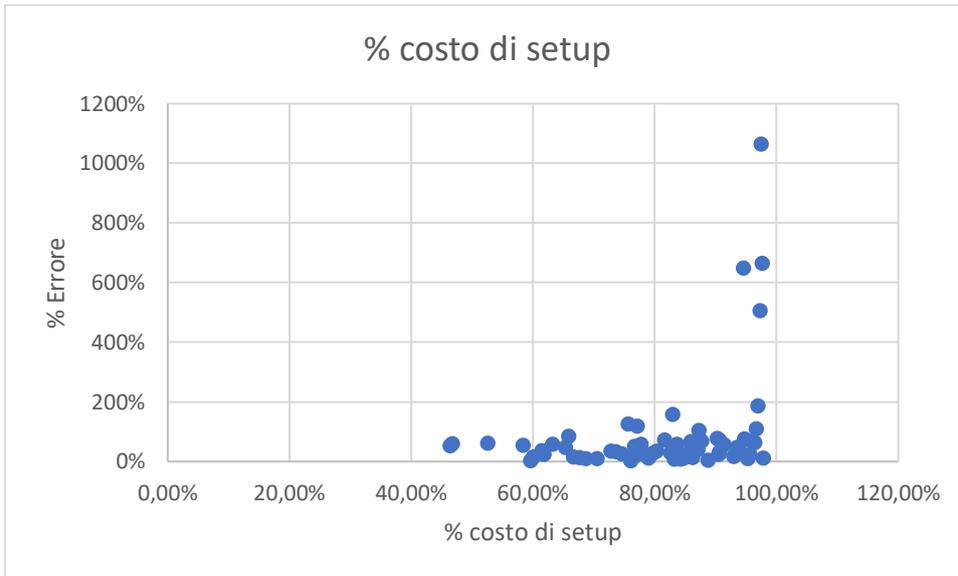


Figura 5.22 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale – LC9

Si possono ripetere le considerazioni fatte per LeanDESIGNER nel capitolo 5.3.2.2.

- N° di utensili stimanti

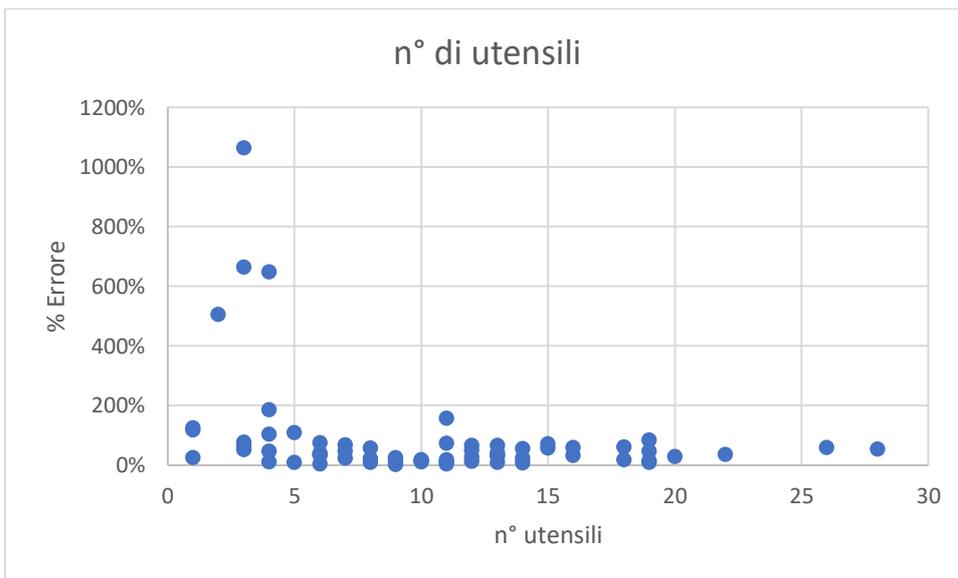


Figura 5.23 - Andamento dell'errore percentuale in funzione del n° di utensili stimati – LC9

Da questo grafico si può vedere come al crescere del numero di utensili il range di errore percentuale decresce.

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7
•	•
•	•
•	•
Codice 28	139
Codice 29	150
Codice 30	22,1
Codice 31	140
Codice 32	135,5
Codice 33	85,5
Codice 34	96
Codice 35	165
Media	58,49

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	75
Codice 2	63
Codice 3	39
Codice 4	140
Codice 5	67,5
Codice 6	145
Codice 7	26,85
•	•
•	•
•	•
Codice 24	245
Codice 25	185
Codice 26	145
Codice 27	148
Codice 28	97
Codice 29	485
Codice 30	850
Codice 31	710
Media	269,98

(B)

Figura 5.24 - Costo medio per pezzi con n° utensili da 1 a 10 (A) e da 11 a 28 (B)

$$\% \text{ Errore_medio}_{1 \rightarrow 10} = 91\%$$

$$\% \text{ Errore_medio}_{11 \rightarrow 28} = 35\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore_ass_medio}_{1 \rightarrow 10} = 91\% * 58,49\text{€} = 53,23\text{€}$$

$$\text{Errore_ass_medio}_{11 \rightarrow 28} = 35\% * 269,98\text{€} = 94,49\text{€}$$

- Classe di appartenenza

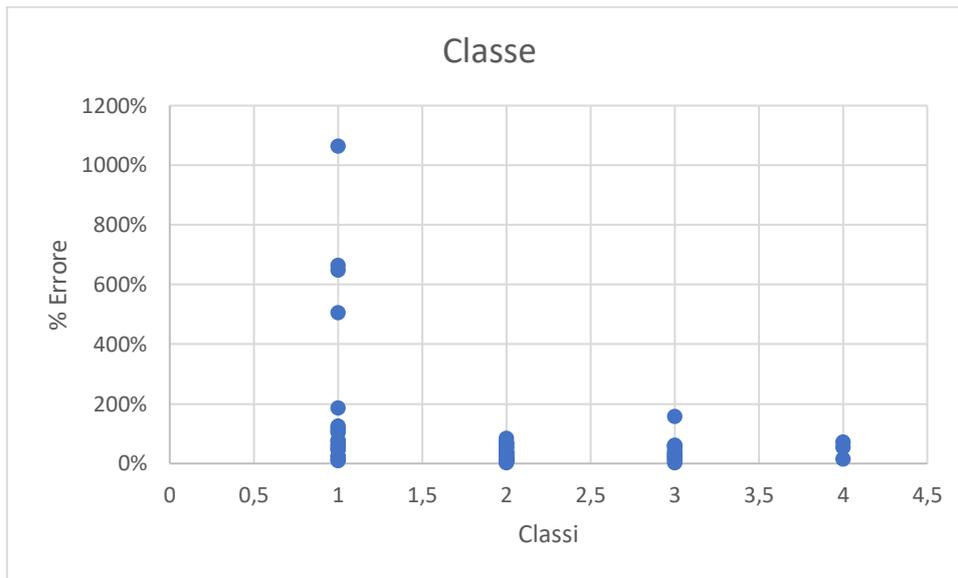


Figura 5.25 - Errore percentuale per le varie classi di produzione – LC9

Gli errori percentuali di ciascuna classe sono:

$$\% \text{ Errore}_{\text{medio}_D} = 207\%$$

$$\% \text{ Errore}_{\text{medio}_C} = 32\%$$

$$\% \text{ Errore}_{\text{medio}_B} = 41\%$$

$$\% \text{ Errore}_{\text{medio}_A} = 46\%$$

Gli errori assoluti medi di ciascuna classe risultano:

$$\text{Errore}_{\text{ass}_{\text{medio}_D}} = 207\% * 15,47\text{€} = 32,02\text{€}$$

$$\text{Errore}_{\text{ass}_{\text{medio}_C}} = 32\% * 114,57\text{€} = 36,66\text{€}$$

$$\text{Errore}_{\text{ass}_{\text{medio}_B}} = 41\% * 347,73\text{€} = 142,57\text{€}$$

$$\text{Errore}_{\text{ass}_{\text{medio}_A}} = 46\% * 681,67\text{€} = 313,57\text{€}$$

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7
Codice 8	21
Codice 9	42
Codice 10	2
Codice 11	13
Codice 12	2
Codice 13	5
Codice 14	25,5
Codice 15	5
Codice 16	15
Codice 17	22
Codice 18	12,5
Codice 19	31,5
Media	15,47

(D)

Codice	Prezzo d'acquisto
Codice 1	580
Codice 2	580
Codice 3	100
Codice 4	760
Codice 5	620
Codice 6	660
Codice 7	235,5
Codice 8	165
Codice 9	245
Codice 10	185
Codice 11	145
Codice 12	148
Codice 13	97
Media	347,73

(B)

Codice	Prezzo d'acquisto
Codice 1	75
Codice 2	100
Codice 3	63
Codice 4	39
Codice 5	250
Codice 6	140
Codice 7	70
Codice 8	67,5
Codice 9	21
Codice 10	145
Codice 11	26,85
Codice 12	156
Codice 13	279,17
Codice 14	160
Codice 15	95
Codice 16	35
Codice 17	141
Codice 18	98
Codice 19	129
Codice 20	139
Codice 21	150
Codice 22	195
Codice 23	22,1
Codice 24	140
Codice 25	190
Codice 26	140
Codice 27	135,5
Codice 28	53,45
Codice 29	85,5
Codice 30	96
Media	114,569

(C)

Codice	Prezzo d'acquisto
Codice 1	485
Codice 2	850
Codice 3	710
Media	681,67

(A)

Figura 5.26 - Costo medio delle varie Classi

Possiamo concludere che per il valore assoluto dell'errore aumenta all'aumentare della complessità del pezzo come detto per LD.

5.3.2.5. LeanCOST ACMA

- Tolleranza più spinta in μm .

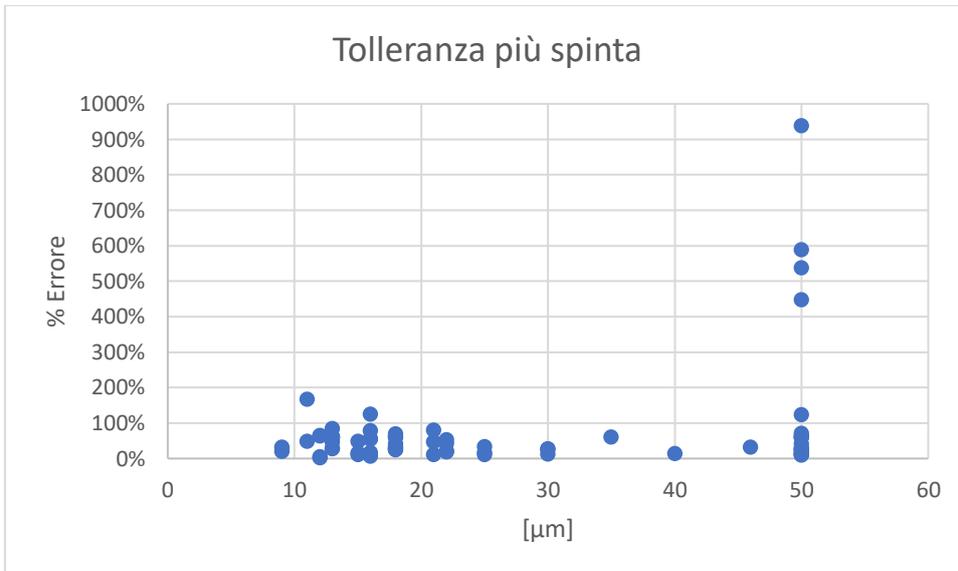


Figura 5.27 - Dati Tolleranza - % Errore – LC ACMA

Si possono ripetere in maniera identica le considerazioni fatte nei capitoli precedenti.

- Presenza di geometrie fuori asse

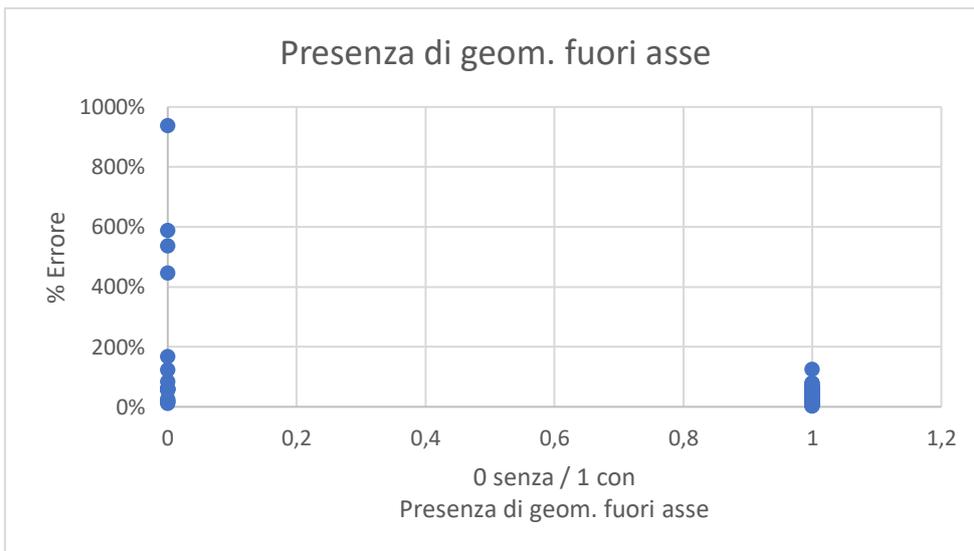


Figura 5.28 - Influenza della presenza di geometrie fuori asse sulla precisione di LC ACMA

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	7
Codice 7	21
Codice 8	42
Codice 9	2
Codice 10	2
Codice 11	5
Codice 12	25,5
Codice 13	5
Codice 14	12,5
Codice 15	31,5
Codice 16	250
Codice 17	35
Media costo	28,94

(A)

Codice	Prezzo d'acquisto [€]
Codice 1	37
Codice 2	13
Codice 3	15
Codice 4	22
Codice 5	75
Codice 6	100
Codice 7	63
•	•
•	•
•	•
Codice 42	245
Codice 43	185
Codice 44	145
Codice 45	148
Codice 46	97
Codice 47	485
Codice 48	850
Codice 49	710
Media	202,54

(B)

Figura 5.29 - Prezzi per pezzi torniti senza presenza di geometrie fuori asse (A) e con geometrie fuori asse (B)

I valori medi dell'errore percentuale sono:

$$\% \text{ Errore}_{medio_0} = 113\%$$

$$\% \text{ Errore}_{medio_1} = 45\%$$

Sfruttando questi dati sul prezzo d'acquisto medio si ottiene un errore assoluto di:

$$\text{Errore}_{ass_medio_0} = 113\% * 28,94€ = 32,70€$$

$$\text{Errore}_{ass_medio_1} = 45\% * 202,54€ = 91,14€$$

- % costo di setup

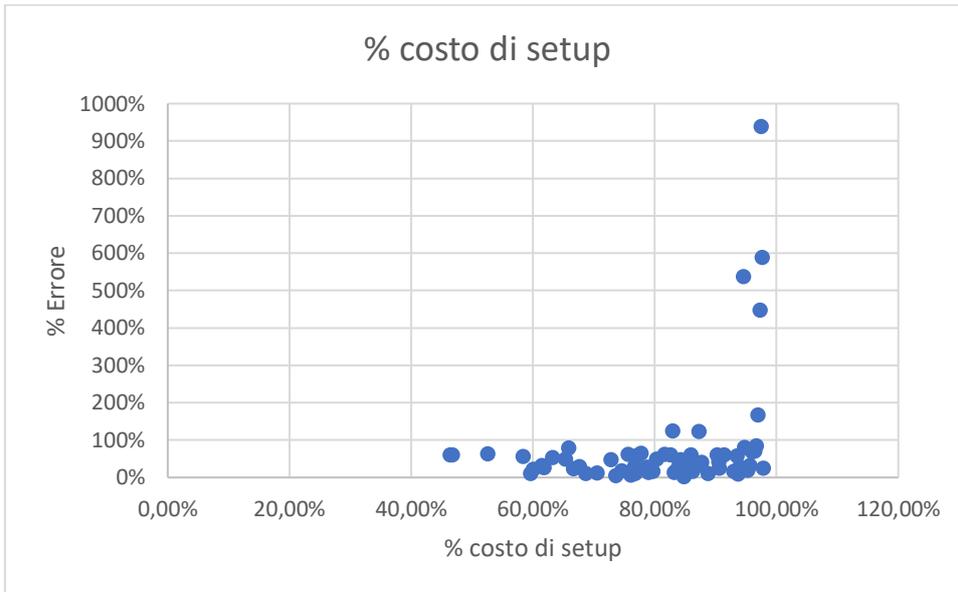


Figura 5.30 - Andamento dell'errore percentuale in funzione della % del costo di setup sul prezzo totale – LC ACMA

Si possono ripetere le considerazioni nel capitolo 5.3.2.2.

- Classe di appartenenza

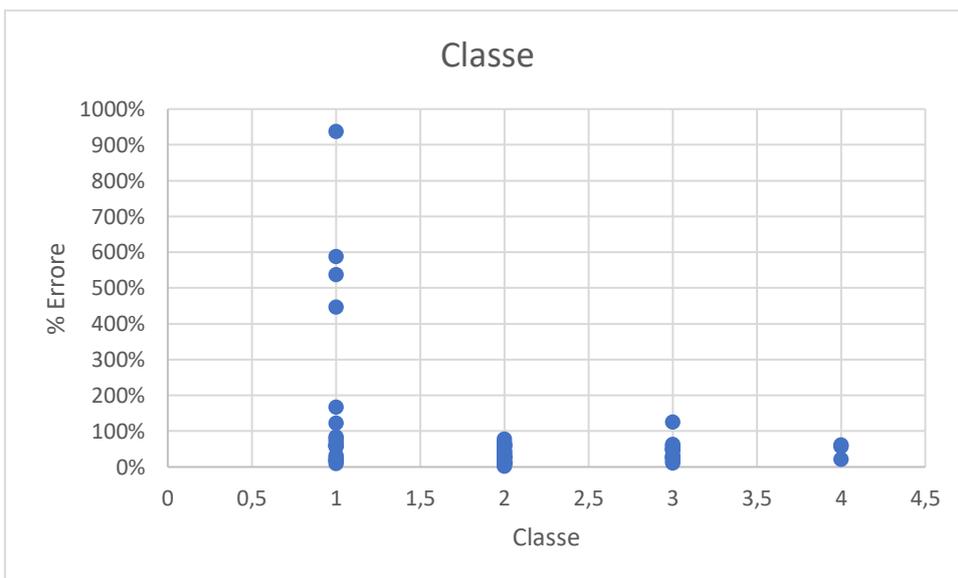


Figura 5.31 - Errore percentuale per le varie classi di produzione – LC ACMA

Gli errori percentuali di ciascuna classe sono:

$$\% \text{ Errore}_{\text{medio}_D} = 178\%$$

$$\% \text{ Errore}_{\text{medio}_C} = 32\%$$

$$\% \text{ Errore}_{\text{medio}_B} = 44\%$$

$$\% \text{ Errore}_{\text{medio}_A} = 46\%$$

Gli errori assoluti medi di ciascuna classe risultano:

$$Errore_{ass_medio_D} = 178\% * 15,47\text{€} = 27,53\text{€}$$

$$Errore_{ass_medio_C} = 32\% * 114,57\text{€} = 36,66\text{€}$$

$$Errore_{ass_medio_B} = 44\% * 347,73\text{€} = 153,00\text{€}$$

$$Errore_{ass_medio_A} = 46\% * 681,67\text{€} = 313,57\text{€}$$

Codice	Prezzo d'acquisto [€]
Codice 1	8
Codice 2	10
Codice 3	6
Codice 4	24
Codice 5	5,5
Codice 6	37
Codice 7	7
Codice 8	21
Codice 9	42
Codice 10	2
Codice 11	13
Codice 12	2
Codice 13	5
Codice 14	25,5
Codice 15	5
Codice 16	15
Codice 17	22
Codice 18	12,5
Codice 19	31,5
Media	15,47

(D)

Codice	Prezzo d'acquisto
Codice 1	580
Codice 2	580
Codice 3	100
Codice 4	760
Codice 5	620
Codice 6	660
Codice 7	235,5
Codice 8	165
Codice 9	245
Codice 10	185
Codice 11	145
Codice 12	148
Codice 13	97
Media	347,73

(B)

Codice	Prezzo d'acquisto
Codice 1	75
Codice 2	100
Codice 3	63
Codice 4	39
Codice 5	250
Codice 6	140
Codice 7	70
Codice 8	67,5
Codice 9	21
Codice 10	145
Codice 11	26,85
Codice 12	156
Codice 13	279,17
Codice 14	160
Codice 15	95
Codice 16	35
Codice 17	141
Codice 18	98
Codice 19	129
Codice 20	139
Codice 21	150
Codice 22	195
Codice 23	22,1
Codice 24	140
Codice 25	190
Codice 26	140
Codice 27	135,5
Codice 28	53,45
Codice 29	85,5
Codice 30	96
Media	114,569

(C)

Codice	Prezzo d'acquisto
Codice 1	485
Codice 2	850
Codice 3	710
Media	681,67

(A)

Figura 5.32 - Costo medio delle varie Classi

Si può concludere che il valore assoluto dell'errore aumenta all'aumentare della complessità del pezzo.

5.4. TEMPISTICHE NECESSARIE PER LA STIMA DEI COSTI

Come spiegato nell'introduzione un aspetto fondamentale riguarda il tempo per effettuare la stima con i vari programmi.

Durante l'analisi dei pezzi torniti infatti si annotano le tempistiche di:

- a. Analisi e trascrizione delle caratteristiche di rugosità, tolleranze e filettature nel CAD 3D: si tiene conto del tempo necessario per riportare le varie rugosità, tolleranze e filettature dalla tavola 2D al disegno CAD 3D;
- b. Analisi tramite software LeanCOST: si tiene conto del tempo impiegato ad analizzare il componente tramite il software LeanCOST facendone le opportune modifiche e correzioni dopo aver importato nel programma il modello CAD 3D;
- c. Analisi tramite software LeanDESIGNER: si tiene conto del tempo impiegato ad analizzare il componente tramite il software LeanDESIGNER facendone le opportune modifiche e correzioni dopo aver importato nel programma il modello CAD 3D.

Ciascuna delle grandezze sopra elencate è stata inserita in una colonna di un foglio excel. Con queste colonne si sono poi calcolati i tempi medi per ciascuna fase sviluppando una semplice media aritmetica. I risultati ottenuti sono riassunti nella Tabella 5.10:

Fase a. [min]	Fase b. [min]	Fase c. [min]
9,80	4,04	7,41

Tabella 5.10 - Analisi tempi

Dall'analisi si evince come la fase più dispendiosa in termini di tempo impiegato sia la fase a.

Essa è fondamentale per le successive quindi “ineliminabile”. L’unico modo per elidere questo tempo morto sarebbe quella di sfruttare la funzione presente nei software LeanCOST e LeanDESIGNER che permette di leggere caratteristiche di rugosità, tolleranza e filettature direttamente dai programmi CAD da cui si importa il modello.

Per fare ciò però sarà necessario che il progettista inserisca, durante la fase di disegno del prodotto, tutti i valori necessari alla fase di preventivazione costi, proprio sul modello 3D (e non solamente in tavola).

6. CASO STUDIO: PARTE 2

6.1. PREPARAZIONE DEL DATABASE

Nella seconda parte del lavoro di tesi l'obiettivo è stato quello di determinare un modello parametrico in grado di predire con una buona accuratezza il costo di future componenti, fornendo in input soltanto determinate caratteristiche di conoscenza del progettista.

Il modello parametrico deve costificare i codici puntando ai costi preventivati da LeanCOST ver.11 che sarà dunque la nuova variabile dipendente.

Per iniziare bisogna preparare il database facendo in modo che sia popolato da informazioni conoscibili dal progettista in fase di design. Si utilizza dunque il database del capitolo 4 a cui si aggiungono alcune voci:

- Trattamento termico
- Trattamento superficiale
- Tipologia finito (solo per i torniti): pieno / cavo
 - o Per i codici cavi si aggiunge il parametro D4 che indica il diametro interno del pezzo.
- Rapporto volume finito/ volume d'ingombro
- Rugosità standard
- Superficie rugosità standard / superficie totale
- Numero di filettature

Si eliminano invece le voci di:

- Classe di appartenenza
- Numero centri di costo
- Numero di utensili

Una volta fatto ciò si va ad aumentare il numero di record del database utilizzando i dieci intervalli di costo calcolati nel capitolo precedente con LeanCOST 11.

In questo modo, dagli iniziali 67 records per la tornitura, se ne ottengono ben 670. Allo stesso modo per i 51 file di fresatura si ottengono 510 records.

6.2. ANALISI CON TOOL SPERIMENTALE

Definito il database si trasformano i valori non numerici come trattamento termico, trattamento superficiale, materiale ecc. in valori numerici. A questo punto si avvia il tool sperimentale che prende in input questi dati ed esegue le seguenti fasi in automatico:

1. **Analisi Pearson:** come nel capitolo precedente si eliminano le caratteristiche a bassa correlazione con la variabile dipendente (che questa volta risulta essere il costo preventivato da LeanCOST ver.11).
2. **Split del database:** si divide in maniera random il database di partenza in training set e testing set. Nel nostro caso le percentuali di splittaggio sono rispettivamente del 80% e 20%.
3. **Analisi di questi database attraverso varie metodologie:** si cerca di determinare un modello utilizzando le tecniche di regressione lineare, rete neurale, random forest e gradient boosting. Esse sono implementate già nel tool.

4. **Confronto tra i risultati ottenuti con le varie metodologie e scelta del modello migliore:** allenati e creati i modelli con le varie tecniche di machine learning proposte, si procede nel testare i vari modelli ottenuti. Si utilizzano i dati di testing ottenuti inizialmente dallo split data del database che si sta processando. Questa fase è cruciale in quanto permette di capire quale tecnica risulta avere le migliori performance in relazione anche al database che viene processato. Infatti, in questa fase si vanno a confrontare, per ogni database, i risultati ottenuti con le differenti tecniche proposte. Il confronto può essere eseguito sulla base di differenti metriche di misura, tutti chiaramente rappresentanti l'errore tra il valore predetto dal modello e quello di riferimento.

In questo studio è stato deciso di valutare la performance del modello attraverso il MAPE.

- MAPE: Mean Absolute Percentage Error:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|y_i - \hat{y}_i|}{y_i} \times 100 \right) \quad (6.1)$$

Inoltre, possono essere utilizzati anche gli errori assoluti:

- MSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6.2)$$

Dove:

- y : valore di actual cost o should cost stimato;
- \hat{y} : valore di actual cost o should cost noto;
- n : numero totale di coppie.

Ogni metodologia avrà una certa qualità del risultato misurata dai parametri R^2 e MAPE. Il programma va ad eseguire l'analisi con ciascuna metodologia e ottiene quattro coppie di valori (R^2 ; MAPE).

Confrontando questi dati, il programma sceglie il modello di maggior qualità.

5. **Export del modello:** scelto il modello ottimale, il programma crea un file .zip che contiene o la formula, o l'albero decisionale o la rete neurale ottenuta dall'analisi. Questo dipende da quale metodo viene giudicato come ottimale dal tool. Tale file sarà utilizzato in uno studio successivo, per implementare nel tool la funzione di preventivazione di nuovi pezzi. Sarà dunque possibile costificare nuovi pezzi non appartenenti al database da cui è stato creato il modello.

6. **Export analisi:** oltre al file .zip del punto precedente il tool va a creare altri tre file Excel:

- Database iniziale completo;
- Database usato per il training;
- Database usato per il testing.

Il foglio Excel che presenta il database di testing può essere definito come il primo utilizzo del modello trovato attraverso il database di training.

Ciò che si trova in questo file è la costificazione attraverso i parametri significativi individuati nella fase 1 dei codici splittati nel database set nella fase 2. Si confronta l'output del modello con il costo del database iniziale. Con questi dati si trova la qualità del modello esplicitata dai parametri R^2 e MAPE.

Riassumendo si può dire che:

INPUT:

- Database iniziale

OUTPUTS:

- un file .zip da utilizzare per creare un'appendice al tool utile per la preventivazione di nuovi pezzi.
- tre file Excel di cui uno, quello contenente il database di testing, che permette di capire la qualità del modello.

6.3. RISULTATI DEL MODELLO OTTIMALE

6.3.1. Tornitura

Per la tornitura i parametri ad alta correlazione dopo l'analisi Pearson sono:

- Tipologia di materiale
- Rivestimento superficiale
- D1
- D3
- Rapporto volume finito / volume ingombro
- Superficie a rugosità minima
- Superficie a rugosità standard
- Presenza di geometrie fuori asse
- Numero di filettature
- Lotto d'acquisto

Tali parametri saranno gli inputs da fornire al tool che verrà sviluppato per costificare i nuovi codici.

In output avremo il "*predicted cost*".

Per verificare la qualità del modello, come anticipato, si va sfruttare il database di test nel quale si vanno ad analizzare i diversi codici ottenendo la colonna dei "*Predicted Cost*". Questa colonna viene poi confrontata con la colonna dei costi predetti da LeanCOST ("*LeanCOST 11 Cost*") per ottenere gli indicatori MAPE e R².

Si ottengono i seguenti risultati:

MAPE	MPE (positivi)	MPE (negativi)
6,74%	6,02%	-7,40%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,02%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
38,45%	37,11%	-38,45%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
7,49%	5,32%	7,09%

Tabella 6.1 - MAPE Modello parametrico – Tornitura

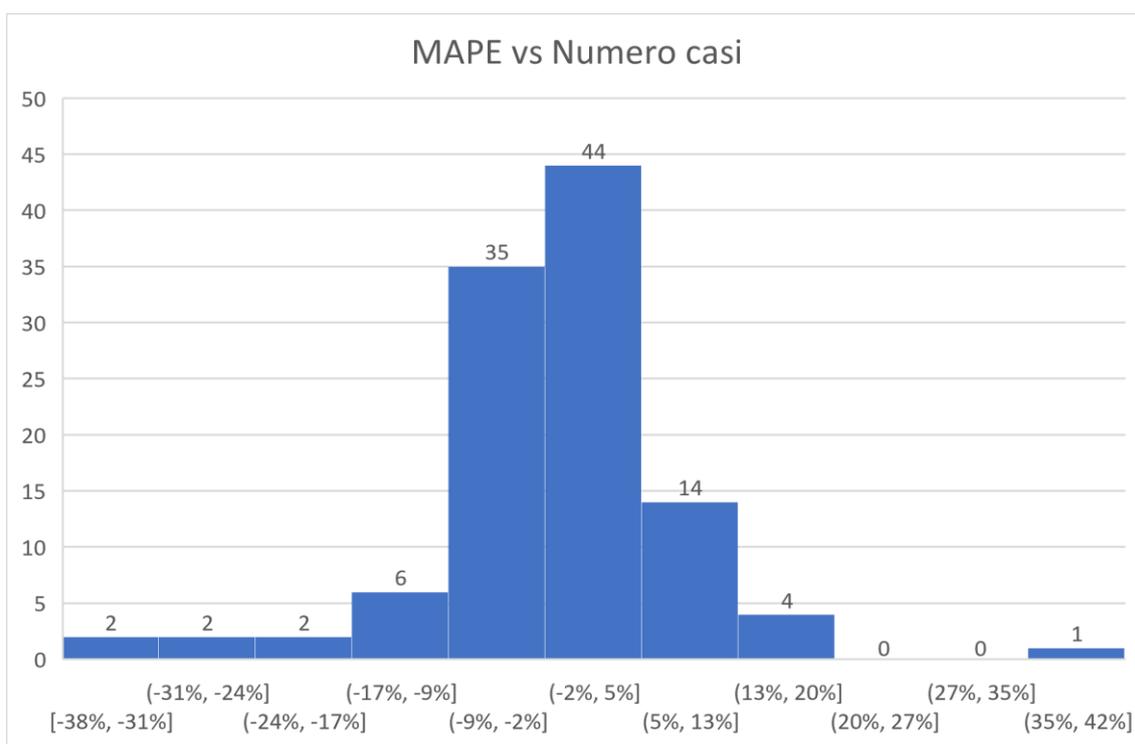


Figura 6.1 - Numero casi a diverso MAPE - Tornitura

R^2
0,990900072

Tabella 6.2 – R^2 modello parametrico - Tornitura

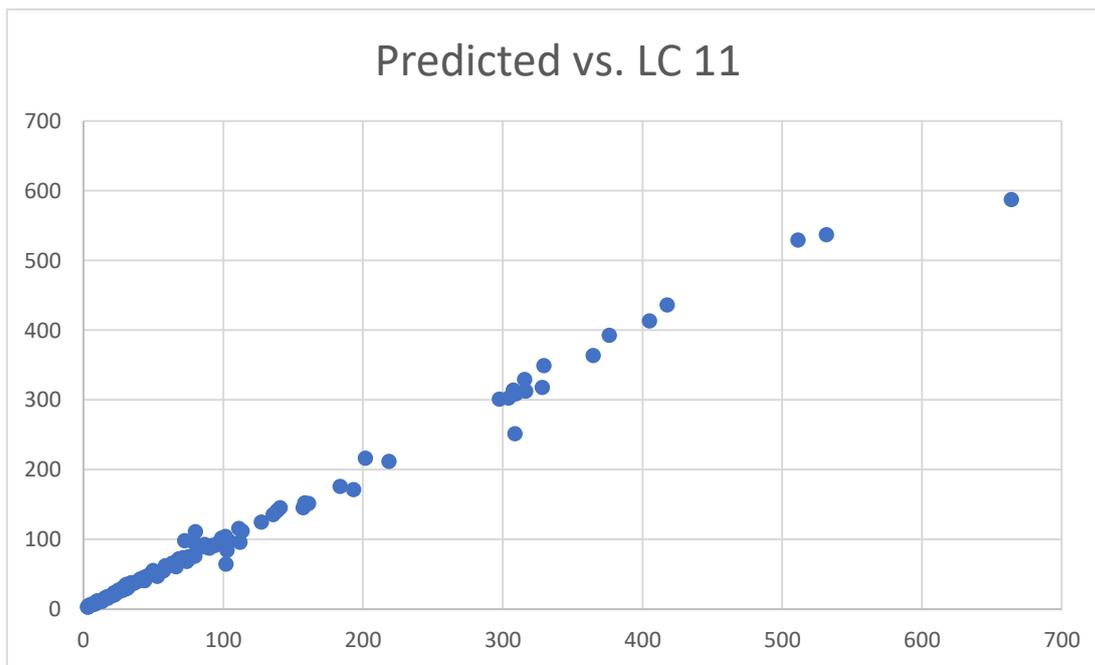


Figura 6.2 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 - Tornitura

Possiamo vedere come il MAPE che rappresenta l'errore percentuale medio della previsione tramite modello parametrico rispetto al costo preventivato da LeanCOST ver.11 sia inferiore al 7%. In particolare, considerando l'istogramma in Figura 6.1 si vede come su 110 analisi, si hanno 79 casi per cui l'errore va da $-9,00\% < MAPE < 5,00\%$.

Nel 84,5%, ovvero 93 volte su 110, si ha: $-9,00\% < MAPE < 13,00\%$.

Se si costruisce un grafico tale da avere sull'asse x il costo preventivato da LeanCOST 11 e sull'asse y il costo predetto dal modello trovato, si trova un andamento simile ad una retta a 45° (che rappresenta la coincidenza tra le due voci di *predicted cost* e *LC Cost*).

Per completare l'analisi si utilizza il metodo *Numeric Outlier-quartile* per individuare eventuali *Outlier*, ovvero quei codici per cui si ha un errore percentuale molto alto.

Il rilevamento dei valori anomali è calcolato mediante il range interquartile.

Nel dettaglio vengono calcolati il primo ed il terzo quartile (Q_1 , Q_3) e si avrà che un valore anomalo è quel valore x_i che si trova al di fuori dell'intervallo interquartile.

È possibile pertanto definirlo attraverso l'espressione:

$$x_i > Q_3 + k(IQR) \vee x_i < Q_1 - k(IQR)$$

Dove $IQR = Q_3 - Q_1$ e $k \geq 0$

Per questo studio è stato preso il valore del moltiplicatore interquartile $k = 1,5$.

Gli outliers sono identificabili anche nell'istogramma in Figura 6.1: sono i valori con alti valori di MAPE di cui si hanno pochi casi.

6.3.1.1. Tornitura con parametri noti dal progettista

Per un progettista non è facile sapere alcuni dei parametri di input necessari al modello descritto nel paragrafo 6.3.1 come, ad esempio, l'estensione delle superfici o i rapporti tra volume del finito e quello del grezzo.

Si indaga dunque sulla possibilità di costruire il modello usando solo quei parametri che sono quasi certamente noti a un ingegnere progettista.

Si considerano come nuovi parametri di input:

- Tipo materiale
- D1
- D3
- Rugosità minima
- Presenza di geometrie fuori asse
- Lotto

Si riavvia la procedura allenando il modello con gli stessi dati ma tenendo in considerazione i nuovi “cost drivers” ottenendo i seguenti parametri di qualità:

MAPE	MPE (positivi)	MPE (negativi)
9,45%	7,62%	-11,27%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,08%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
62,57%	32,75%	-62,57%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
10,35%	6,22%	10,57%

Tabella 6.3 - Numero casi a diverso MAPE - Tornitura v2

R ²
0,989267661

Tabella 6.4 – R² modello parametrico - Tornitura v2

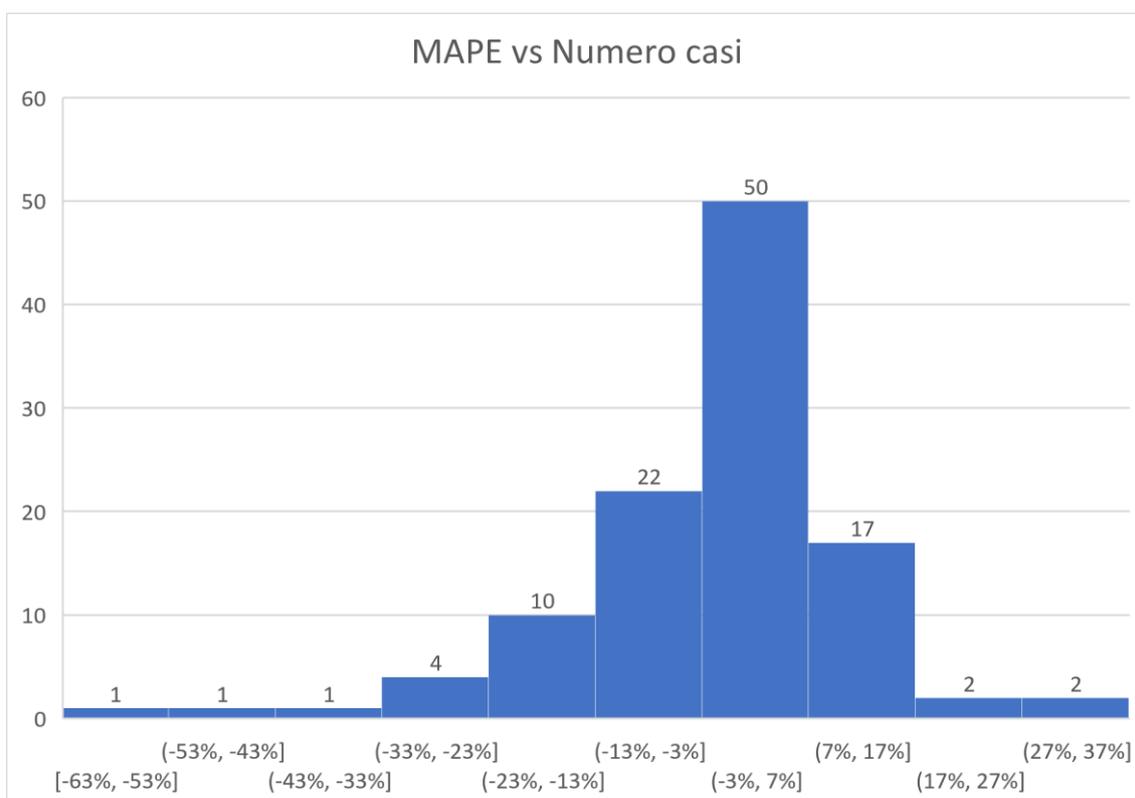


Figura 6.3 - Numero casi a diverso MAPE - Tornitura v2

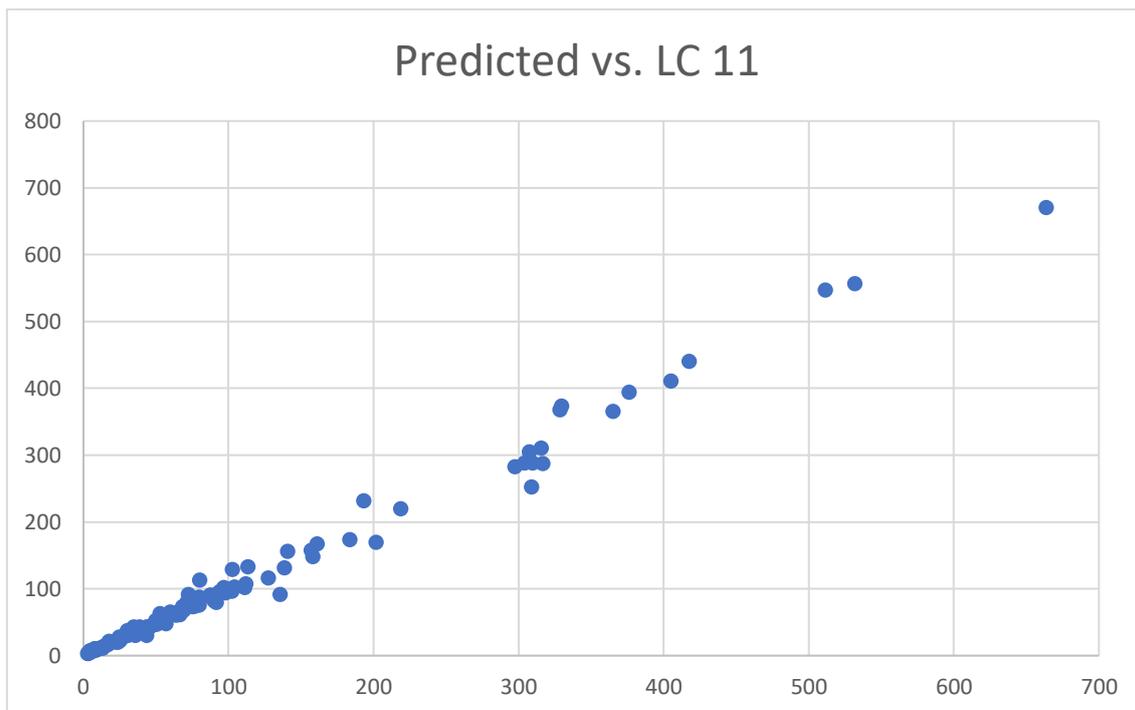


Figura 6.4 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 - Tornitura v2

L'andamento dell'istogramma permette di vedere la qualità del modello trovato dal tool (Figura 6.3). In particolare, considerando l'istogramma si vede come su 110 analisi, si hanno 89 casi per cui l'errore va da $-13,00\% < MAPE < 17,00\%$ (nel 80,9% dei casi).

Il valore di MAPE medio risulta essere del 9,45%.

6.3.2. Fresatura

Per la fresatura i parametri ad alta correlazione dopo l'analisi Pearson sono:

- Tipologia di materiale
- Tipologia grezzo
- D1
- D2
- D3
- Numero fori
- Superficie a rugosità standard
- Lotto d'acquisto

Come nel caso precedente, fornendo in questi dati al tool si otterrà in output il "Predicted Cost".

In maniera identica alla tornitura si calcola la qualità del modello ottenuto dal database iniziale calcolando gli indici di MAPE e R².

Si ottengono i seguenti risultati:

MAPE	MPE (positivi)	MPE (negativi)
8,14%	6,33%	-10,48%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,03%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
97,04%	32,49%	-97,04%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
12,74%	6,28%	12,49%

Tabella 6.5 - MAPE Modello parametrico – Fresatura

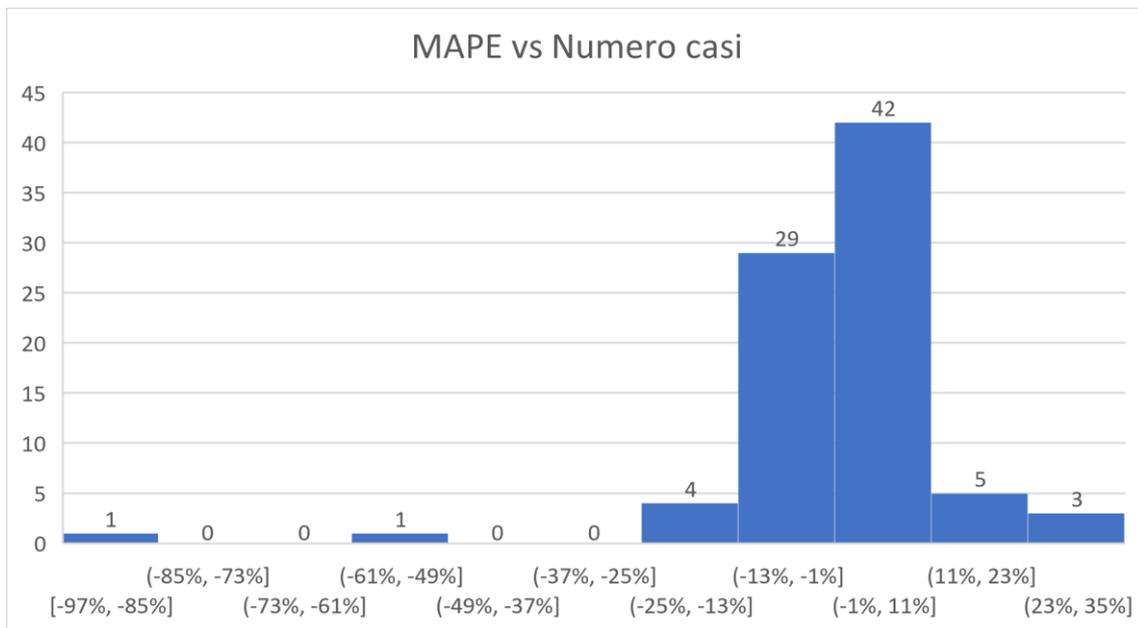


Figura 6.5 - Numero casi a diverso MAPE – Fresatura

R ²
0,951595603

Tabella 6.6 – R² modello parametrico - Fresatura

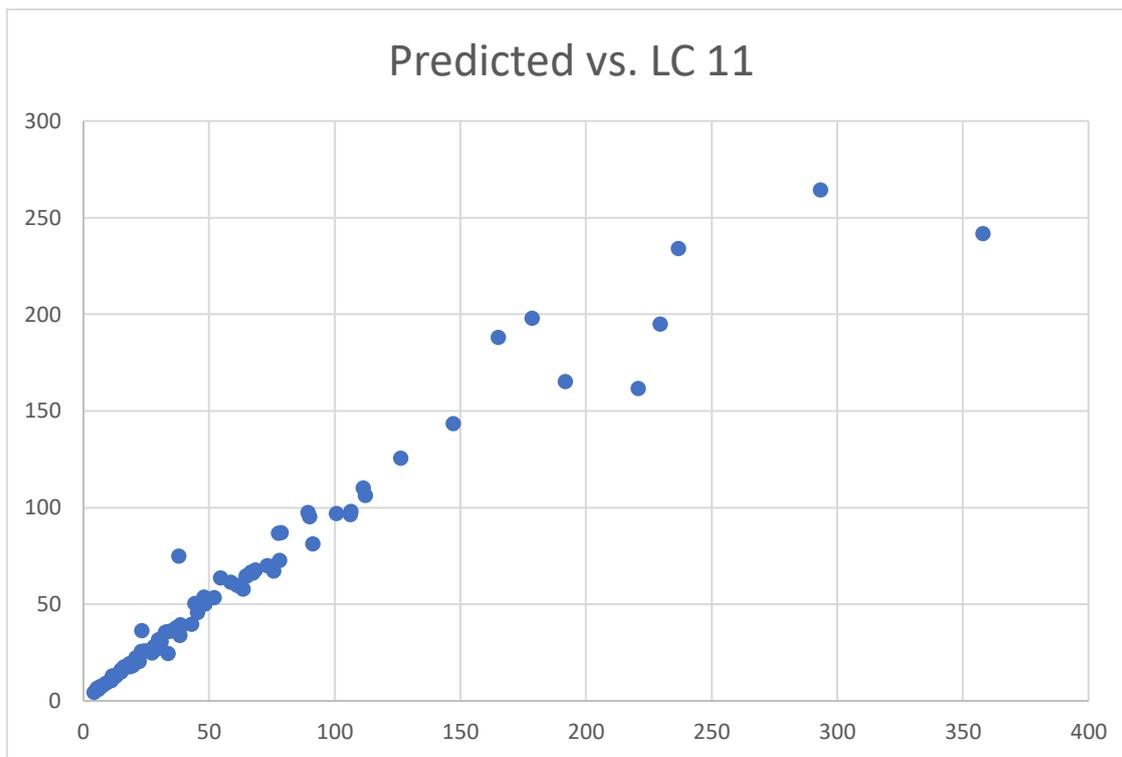


Figura 6.6 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 - Fresatura

L'andamento dell'istogramma permette di vedere la qualità del modello trovato dal tool. Esso rappresenta il numero di casi a vari valori di MAPE (Figura 6.5). In particolare, considerando l'istogramma si vede come su 85 analisi, si hanno 71 casi per cui l'errore va da $-13,00\% < MAPE < 11,00\%$ (nel 83,5% dei casi).

Se si costruisce un grafico tale da avere sull'asse x il costo preventivato da LeanCOST 11 e sull'asse y il costo predetto dal modello trovato, si trova un andamento simile ad una retta a 45° (che rappresenta la coincidenza tra le due voci di *predicted cost* e *LC Cost*). I dati che si discostano di più sono quei dati che hanno un costo preventivato alto quindi per questi codici sarà necessaria un'analisi più approfondita.

Il valore di MAPE medio risulta essere del 8,14%.

6.3.2.1. Fresatura con parametri noti dal progettista

Come nel caso dei torniti, si indaga sulla possibilità di costruire un modello su quei parametri che sono quasi certamente noti a un ingegnere progettista.

Si considerano come nuovi parametri di input:

- Tipo materiale
- D1
- D2
- D3
- Numero Fori
- Volume grezzo
- Tolleranza più spinta
- Finitura più spinta
- Lotto

Si riavvia la procedura allenando il modello con gli stessi dati ma tenendo in considerazione i nuovi “cost drivers” ottenendo i seguenti parametri di qualità:

MAPE	MPE (positivi)	MPE (negativi)
10,90%	10,90%	-10,91%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,29%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
62,49%	48,91%	-62,49%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
11,33%	7,68%	10,75%

Tabella 6.7 - MAPE Modello parametrico – Fresatura v2

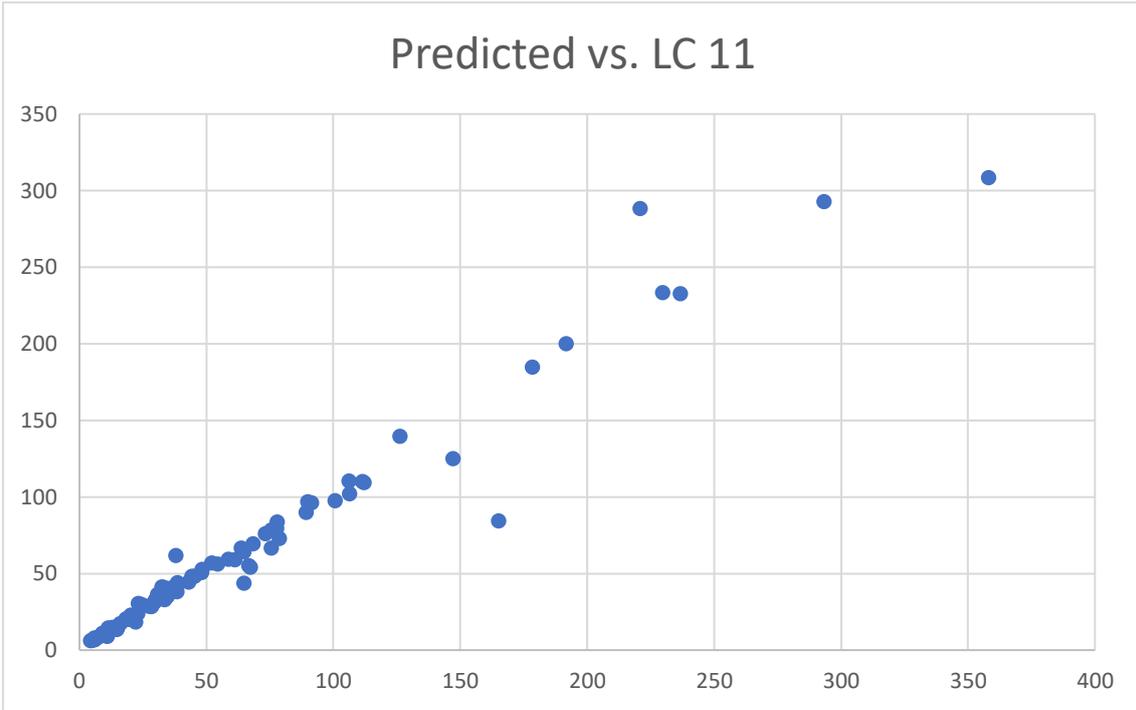


Figura 6.7 - Costo previsto parametricamente - Costo preventivato con LeanCOST ver.11 – Fresatura v2

R^2
0,956047134

Tabella 6.8 - R^2 modello parametrico – Fresatura v2

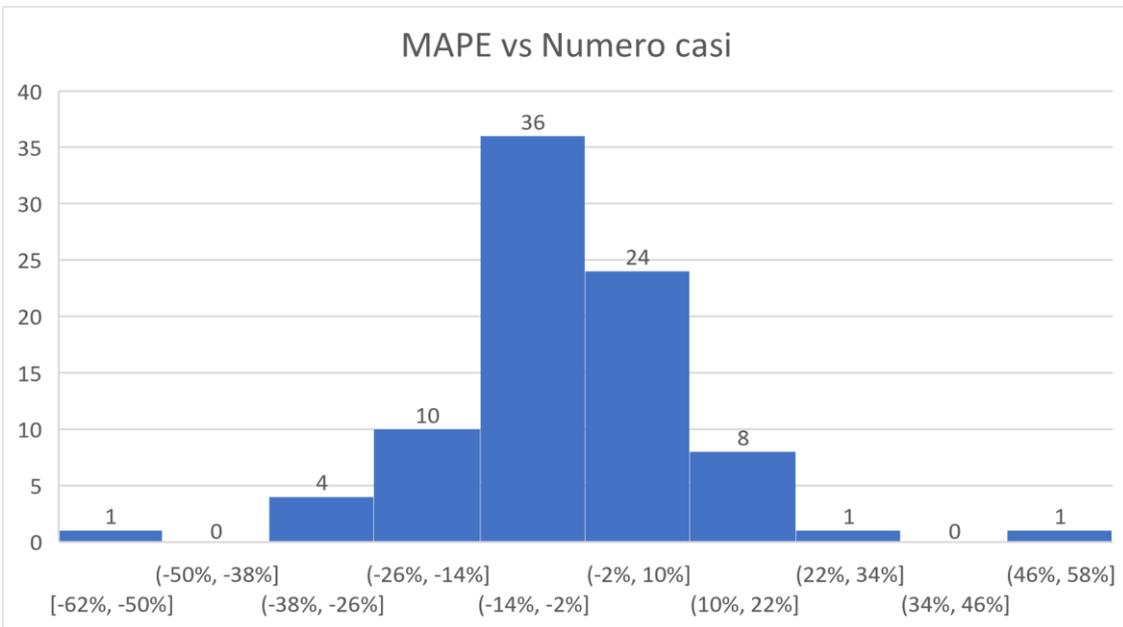


Figura 6.8 - Numero casi a diverso MAPE - Fresatura v2

L'andamento dell'istogramma permette di vedere la qualità del modello trovato dal tool (Figura 6.8). In particolare, considerando l'istogramma si vede come su 85 analisi, si hanno 60 casi per cui l'errore va da $-14,00\% < MAPE < 10,00\%$ (nel 70,6% dei casi).

Il valore di MAPE medio risulta essere del 10,90%.

6.3.3. Convalida del modello

Per testare la bontà del modello trovato si vanno ad esaminare 8 nuovi codici: 4 per la famiglia dei torniti e 4 per la famiglia dei fresati. Si fa questo affrontando una prima preventivazione attraverso LeanCOST 11 la quale rappresenterà il costo obiettivo. Si vanno successivamente ad utilizzare i modelli parametrici ottenuti e si confrontano quest'ultime stime con quella obiettivo del software analitico.

In particolare, si sviluppano due modelli parametrici per ciascuna famiglia tecnologica di cui il modello ottimale viene utilizzato in due modi:

- Approssimativa: i parametri in input sono inseriti con calcoli di massima da parte dell'utilizzatore;
- Precisa: si inseriscono tutti i parametri in maniera puntuale leggendo le varie caratteristiche dal software CAD.

Questa distinzione vuole mostrare l'evoluzione della qualità del modello, dalla fase di "conceptual design" (utilizzo approssimativo), a quella in cui si ha il CAD completamente definito.

I risultati ottenuti sono quelli visibili nelle Tabella 6.9 - Analisi qualità modello parametrico – Torniti e Tabella 6.10:

Torniti	Mod.1	% errore	Mod.1 Preciso	% errore	Mod.2 Preciso	% errore	LC11
codice 1	85,7	26%	80,4	19%	70,5	4%	67,8
codice 2	585,2	-5%	582,7	-6%	694,8	12%	618,5

codice 3	70,2	-31%	69,1	-32%	43,9	-57%	102,3
codice 4	249,6	-24%	286,8	-12%	240,5	-27%	327,4
media		22%		17%		25%	

Tabella 6.9 - Analisi qualità modello parametrico – Torniti

Fresati	Mod.1	% errore	Mod.1 Preciso	% errore	Mod.2 preciso	% errore	LC11
codice 1	84,0	-13%	81,8	-15%	80,2	-17%	96,4
codice 2	99,4	92%	99,4	92%	64,4	24%	51,9
codice 3	253,9	124%	261,6	131%	181,9	61%	113,4
codice 4	262,5	20%	262,5	20%	369,9	69%	218,5
media		62%		64%		43%	

Tabella 6.10 - Analisi qualità modello parametrico – Fresatura

Dai risultati si vede come per i torniti la qualità del modello per i torniti rimane circa costante (intorno al 22%) andando ad essere meno precisa al diminuire della precisione e quantità dei parametri di inputs.

Per i fresati invece, l'errore risulta essere molto più alto del previsto, sia in caso di inputs precisi che nel caso di inputs noti al progettista in fase di conceptual design (errore superiore al 50%).

Questi risultati derivano dalla diversa complessità che abbiamo tra i pezzi torniti e i pezzi fresati. In generale, la prima famiglia risulta essere più prevedibile rispetto a quella dei fresati che mostra una complessità media maggiore.

Inoltre, l'aumentare dell'errore del modello parametrico può essere fittizio facendo l'analisi su sole 4 componenti per famiglia. Alcuni dei pezzi torniti infatti avevano una rugosità generale di 3,2 e di 6,3 mentre il modello è stato allenato su pezzi che avevano rugosità 12,5.

Per avere dati ancor più significativi bisognerebbe dare in pasto un'ampia lista di codici all'algoritmo per poi trarne delle conclusioni più precise.

7. CONCLUSIONI

Lo scopo dello studio è duplice: da una parte ci si impegna a determinare la precisione della stima dei software di costificazione utilizzati (in base ai valori reali di mercato), dall'altra ci si interroga sulla bontà di un eventuale modello parametrico basato sulle tecnologie di machine learning e sulla fattibilità dello stesso.

Si studia dunque il processo di preventivazione basato su software che sfruttano tecniche analitiche di costificazione (Analisi Should Cost). Ottenuti i valori stimati si sono confrontati con i reali valori di mercato per ottenere alcuni parametri come differenza tra costo stimato e costo reale, differenza percentuale tra costo stimato e costo di mercato e valore assoluto di quest'ultimo che sono stati poi utilizzati per creare degli indici statisticamente significativi all'analisi condotta: delta costo totale, media pesata degli errori e media pesata in valore assoluto.

Una volta fatto ciò si è cercato di legare questi indici a delle caratteristiche proprie del pezzo da produrre. In questo modo è stato possibile capire se ci fossero dei particolari materiali, famiglie tecnologiche, forme ecc. che inducessero i programmi di costificazione ad effettuare una preventivazione sbagliata. Bisogna ricordare che ci si sta interfacciando con fornitori esterni, pertanto, molti fattori non sono calcolabili dai programmi come già spiegato nel paragrafo 4.1.

Per la famiglia tecnologica della tornitura, ad eccezione di LD base, si sono individuati dai quattro ai cinque parametri strettamente correlati all'errore percentuale sul pezzo, Tabella 7.1.

LD base	-----	-----	-----	-----	-----
LD	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	n° utensili LC	Classe
LC 11	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	n° utensili LC	Classe
LC 9	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	n° utensili LC	Classe
LC ACMA	Tolleranza più spinta [μm]	Presenza di geometrie fuori asse	% costo di setup	Classe	

Tabella 7.1 - Parametri ad alta correlazione con % errore

Dopo aver determinato attraverso l'analisi Pearson quali sono questi parametri, si studia "come" tali caratteristiche influenzino l'errore di preventivazione.

Si è giunti ad una duplice conclusione.

Gli errori percentuali più elevati si hanno per pezzi di ridotta complessità:

- Basso numero di utensili
- % costo di setup alta
- Classe D
- Assenza di geometrie fuori asse

Si può dire che sicuramente è più semplice commettere un errore percentuale alto su questi pezzi ma bisognerà vedere in che modo andare a adattare il database sul quale si basa il software per ridurre questo errore.

Altra considerazione che si può fare è che in generale, maggiore è la complessità del pezzo, maggiore sarà l'errore assoluto associato allo stesso. Sicuramente sbagliare anche di pochi punti percentuali la costificazione di questi pezzi, va ad influenzare in maniera molto più marcata il budget stimato e quindi grava di più ai fini del risultato aziendale. Questo discostamento è dovuto anche al rischio che deve assumere il fornitore per la produzione di un pezzo di difficile realizzazione. Costo che ricadrà inevitabilmente sull'acquirente.

Per quanto riguarda invece la famiglia tecnologica della fresatura, non si ottiene alcun parametro ad alta correlazione dall'analisi Pearson.

Questo risultato si può interpretare dicendo che per questa famiglia di pezzi, l'errore sarà in genere più "casuale" rispetto ai codici di tornitura.

Concluso questo primo obiettivo, si sposta l'asticella ancor più in alto e ci si interroga sulla possibilità di utilizzare un approccio basato sulle tecnologie di machine learning per ottenere un modello parametrico in grado di determinare il costo di un manufatto ricevendo in input dei semplici parametri conosciuti dal progettista.

Per la tornitura, sfruttando le tecniche di regressione lineare, rete neurale, random forest e gradient boosting implementate all'interno di un software sperimentale, si è giunti alla determinazione di un modello parametrico ad alte prestazioni.

MAPE	MPE (positivi)	MPE (negativi)
6,74%	6,02%	-7,40%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,02%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
38,45%	37,11%	-38,45%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
7,49%	5,32%	7,09%

Tabella 7.2 - Risultati modello parametrico – Tornitura

I dati ottenuti risultano ottimi ma i parametri di input sui quali si basa il modello non sono di facile conoscenza da parte dell'utilizzatore del software.

Si costruisce allora un nuovo modello basato su parametri quasi certamente conosciuti da parte dei progettisti, ottenendo quanto segue:

MAPE	MPE (positivi)	MPE (negativi)
9,45%	7,62%	-11,27%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,08%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
62,57%	32,75%	-62,57%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
10,35%	6,22%	10,57%

Tabella 7.3 - Risultati modello parametrico – Tornitura v2

Per quanto riguarda invece la fresatura si è giunti, attraverso lo stesso procedimento, alla determinazione di un modello parametrico con i seguenti indici qualitativi:

MAPE	MPE (positivi)	MPE (negativi)
8,14%	6,33%	-10,48%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,03%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
97,04%	32,49%	-97,04%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
12,74%	6,28%	12,49%

Tabella 7.4 - Risultati modello parametrico - Fresatura

Anche in questo caso risulta un ottimo modello ma i parametri di input non sono di facile conoscenza da parte del progettista.

Si costruisce allora un altro modello basato su dati noti all'ingegnere:

MAPE	MPE (positivi)	MPE (negativi)
10,90%	10,90%	10,91%
MAPE min	MPE (positivi) min	MPE (negativi) max
0,29%	0,00%	0,00%
MAPE max	MPE (positivi) max	MPE (negativi) min
62,49%	48,91%	62,49%
MAPE dev standard	MPE (positivi) dev standard	MPE (negativi) dev standard
11,33%	7,68%	10,75%

Tabella 7.5 - Risultati modello parametrico - Fresatura v2

I modelli v2 risultano sì, essere peggiori in termini assoluti rispetto ai modelli originali, ma bisogna considerare che questi possono essere utilizzati in fasi anteriori a quella di costificazione del prodotto finito: conceptual design.

Per i torniti l'errore medio ottenuto durante la convalida del modello si attesta intorno al 20% che sale al 25% se riduciamo i parametri di input ai soli valori conoscibili dal progettista.

Per i fresati l'errore risulta essere più alto sia per un utilizzo preciso (circa il 60%) sia per un utilizzo in fase concettuale (circa il 40%).

Questi risultati potrebbero essere dovuti alla diversa complessità tra i pezzi fresati e i pezzi torniti. In generale, il processo di fresatura risulta essere molto meno prevedibile rispetto a quella di tornitura sarà pertanto necessario un numero maggiore di parametri per descrivere in maniera ottimale il codice da analizzare.

Per avere dati ancor più significativi bisognerebbe utilizzare il modello ottenuto su una lista di codici più estesa, per poi trarne delle conclusioni più precise.

Uno sviluppo futuro potrebbe essere quello di creare con i risultati ottenuti un tool che permetta un monitoraggio "live" del prezzo delle componenti.

APPENDICE A

Codice	Codice 2	Codice 1	Ecc.
Famiglia tecnologica	yy	zz	xx
Famiglia materiale	ww	ee	gg
D ₁	tt	yy	uu
D ₂	oo	pp	aa
D ₃
D ₁ /D ₂
D ₂ /D ₃
D ₁ /D ₃
Tipologia grezzo
Volume ingombro
Volume del finito
Rapporto Vol. ing/ Vol. fin.
Tolleranza più spinta in μm
Finitura superficiale più spinta
Estensione superficie e bassa rugosità
Finitura superficiale standard
Estensione superficie a rugosità standard
Rapporto superficie a rugosità bassa su superficie a rugosità standard
% costo di setup
Numeri utensili necessari
Numeri di centri di costo stimati

8. BIBLIOGRAFIA

- [1] A. Niazi, J. S. Dai, "Product cost estimation: technique classification and methodology review."
- [2] NASA Cost Estimating Handbook Version 4.0, Appendix C: Cost Estimating Methodologies.
- [3] Marco Pizzi "Tecniche di parametric cost modelling per component di turbo macchine"
- [4] Sergio Cavalieri, Paolo Maccarrone, Roberto Pinto (2004). «Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry».
- [5] Lorenzo Govoni. «L'Overfitting e l'Underfitting nel machine learning». Available Online: <https://lorenzogovoni.com/overfitting-e-underfitting-machine-learning/>.
- [6] Lorenzo Govoni «Algoritmo Gradient boosting per problemi di regressione». Available online: <https://lorenzogovoni.com/algoritmo-gradient-boosting-per-problemi-di-regressione/>
- [7] Antonio C. Caputo, Pacifico M. Pelagagge (2008)., «Parametric and neural methods for cost estimation of process vessels».
- [8] Gregory Puckett (2016). «Development of a parametric cost estimating model for University of Alaska system renovation construction projects».
- [9] Azionaria Costruzioni Macchine Automatiche *A.C.M.A. S.P.A. su reportaziende.it
- [10] Sala Borsa di Bologna, la nascita di A.C.M.A., su bibliotecasalaborsa.it
- [11] A.C.M.A. Azionaria Costruzioni Macchine Automatiche, Wikipedia, enciclopedia libera.

- [12] Rockwell, R. C. 1975. "Assessment of multicollinearity: The Haitovsky test of the determinant." *Sociological Methods Res.* 3 (3): 308–320. <https://doi.org/10.1177/004912417500300304>
- [13] O. Duran et al. 2012 "Comparisons between two types of neural networks for manufacturing cost estimation of piping elements". *Expert Systems with Applications* 39 (2012) 7788–7795
- [14] Tsung-Han Yeh & S. Deng (2012) Application of machine learning methods to cost estimation of product life cycle, *International Journal of Computer Integrated Manufacturing*,
- [15] Garza, J. and Rouhana, K., 1995. Neural networks versus parameter-based application in cost estimating. *Cost Engineering*,
- [16] Brass, A. M. Gerrard and D. Peel, "Estimating vessel costs via neural networks," *Proceedings of the 13th International Cost Engineering Congress*, London, 1994.
- [17] Tayefeh Hashemi, S., Ebadati, O.M. & Kaur, H. Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. *SN Appl. Sci.* 2, 1703 (2020)
- [18] Tarek Hegazy, Amr Ayed (1998). «Neural network model for parametric cost estimation of highway projects».
- [19] Marjan Čeh, Milan Kilibarda, Anka Lisec, Branislav Bajac (2018). «Estimating the Performance of Random Forest versus Multiple Regression for Predicting Prices of the Apartments».
- [20] Orady E.A., Newcombe R.W., Yellowley I., 1979. Computer AIDED Estimation of the Manufacturing Time for Turned Components. *Proceedings of the Nineteenth International Machine Tool Design and Research Conference*: 295-303. https://doi.org/10.1007/978-1-349-81412-1_36

- [21] Siadat A., Dantan J.Y., Mauchand M., Martin P., 2007. A COST ESTIMATION SYSTEM FOR MANUFACTURING PRODUCT USING ONTOLOGY AND EXEPERT SYSTEM. IFAC Proceedings Volumes 40(18): 37-42. <https://doi.org/10.3182/20070927-4-RO-3905.00008>
- [22] Bouaziz Z., Ben Younes J., Zghal A., 2006. Cost estimation system of dies manufacturing based on the complex machining features. International Journal of Advanced Manufacturing Technololy 28: 262–271. <https://doi.org/10.1007/s00170-004-2179-3>
- [23] Ou-Yang C., Lin T.S., 1997. Developing an Integrated Framework for Feature-Based Early Manufacturing Cost Estimation. International Journal of Advanced Manufacturing Technology 13: 618-629. <https://doi.org/10.1007/BF01350820>
- [24] Jung J.Y., 2002. Manufacturing cost estimation for machined parts based on manufacturing features. Journal of Intelligent Manufacturing 13(4): 227–238. <https://doi.org/10.1023/A:1016092808320>
- [25] Sajadfar N., Ma Y., 2015. A hybrid cost estimation framework based on feature-oriented data mining approach. Advanced Engineering Informatics 29(3): 633-647. <https://doi.org/10.1016/j.aei.2015.06.001>
- [26] Naranje V., Kumar S., Hussein H.M.A, 2014. A Knowledge Based System for Cost Estimation of Deep Drawn Parts. Procedia Engineering 97: 2313-2322. <https://doi.org/10.1016/j.proeng.2014.12.476>
- [27] Chang K.H., 2016. e-Design Computer-Aided Engineering Design, 1st Edition. Academic Press. <https://doi.org/10.1016/C2009-0-63076-2>
- [28] <http://www.get-a-quote.net/>
- [29] <https://galorath.com/seer-for-manufacturing/>
- [30] <https://www.apriori.com/>
- [31] <https://www.custompartnet.com/>

- [32] <https://www.dfma.com/>
- [33] <https://www.microest.com/index.html>
- [34] <https://www.mtisystems.com/index.html>
- [35] <https://www.solidworks.com/it>