



UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACOLTA DI INGEGNERIA

Corso di Laurea Triennale in Ingegneria Meccanica

Dipartimento di Ingegneria Industriale e Scienze Matematiche

*Sviluppo di un modello di impatto ambientale parametrico
per componenti assialsimmetrici basato su tecniche di
Machine Learning*

*Development of a parametric environmental impact model for
axisymmetric components based on Machine Learning techniques*

Relatore:

Prof. Ing. Michele Germani

Tesi di Laurea di:

Monachesi Massimiliano

Anno accademico 2021/2022

[Digitare qui]

[Digitare qui]

Sommario

I. INDICE FIGURE.....	5
II. INDICE TABELLE	6
1 INTRODUZIONE.....	7
2 STATO DELL'ARTE	8
2.1 Sviluppo Sostenibile Nell'Industria 4.0.....	8
2.2 Life Cycle Assessment	10
2.3 Machine Learning Per La Sostenibilità Ambientale.....	14
2.4 Casi Studio.....	15
2.4.1 Costruzione Di Un Sistema Di Trasporto Intelligente a Basse Emissioni Di Carbonio Basato Su BD e ML.....	15
2.4.2 Approccio Basato Sui Dati Per La Scoperta Dei Potenziali Di Risparmio Energetico Nella Fabbrica Di Produzione	16
3 METODO.....	18
3.1 Tecniche Di Stima	18
3.1.1 Tecniche Di Stima Intuitiva	18
3.1.2 Tecnica Di Stima Analogica.....	19
3.1.4 Tecnica Di Stima Parametrica	20
3.2 Metodo Per Lo Sviluppo Di Un Modello D' Impatto Ambientale.....	20
3.3 Strumenti Utilizzati	21
3.3.1 SimaPro	21
3.3.2 RapidMiner.....	22
3.4 Creazione Database	24
.....	24
3.5 Analisi Di Pearson.....	25
3.6 Analisi Machine Learning	28
4 CASO STUDIO.....	30
4.1 Componenti Del Caso Studio	30

[Digitare qui]

4.2 Analisi Del Impatto Ambientale Tramite SimaPro	32
4.3 Raccolta Dati	33
4.4 Calcolo Emissione Di CO2.....	35
4.5 Generazione Del Database Esteso	36
4.5.1 Scalatura Dei Modelli.....	36
4.5.2 Inserimento Del Materiale	37
4.6 Analisi In RapidMiner.....	37
4.6.1 Panoramica Generale.....	38
4.7 Gradient Boosted Trees	41
4.8 Deep Learning	45
4.8.1 Confronto Scelta Dei Parametri.....	48
4.9 Esempi Applicativi	49
5 CONCLUSIONE.....	53
6. BIBLIOGRAFIA	55

[Digitare qui]

I. INDICE FIGURE

Figura 3 . 1 Peso parametro rispetto all'output	23
Figura 3 . 2 Prediction Chart	23
Figura 3 . 3 Grafico di correlazione lineare.....	26
Figura 3 . 4 Correlazione dell'analisi di Pearson	27
Figura 4 . 1 Support Vector Machine.....	39
Figura 4 . 2 Errore relativo delle diverse analisi	40
Figura 4 . 3 Tempo di esecuzione delle diverse analisi.....	40
Figura 4 . 4 Decision Trees.....	42
Figura 4 . 5 Gradient Boosted Trees Weights.....	43
Figura 4 . 6 Predictions Chart Gradient Boosted Trees	44
Figura 4 . 7 Deep Learning Weights.....	46
Figura 4 . 8 Deep Learning Predictions Chart	47
Figura 4 . 9 Deep Learning Weights.....	48
Figura 4 . 10 Componente T001670	49
Figura 4 . 11 Componente TOO42708	50
Figura 4 . 12 Componente "Albero (9)"	50

[Digitare qui]

II. INDICE TABELLE

Tabella 1 Esempio Database.....	24
Tabella 2 Impatto Ambientale dei Materiali	32
Tabella 3 Impatto Ambientale delle Lavorazioni.....	32
Tabella 4 Errore delle Predizioni di Impatto Ambientale	51

1 INTRODUZIONE

Negli ultimi anni, lo sfruttamento di risorse ambientali per fini produttivi è cresciuto in maniera esponenziale, accentuato dai processi di industrializzazione dovuti al continuo aumento della popolazione e della globalizzazione, causando problematiche a livello ambientale e sociale.

Per evitare ciò, al giorno d'oggi risulta sempre più importante per un'azienda conoscere quale sarà l'impatto ambientale di un prodotto in fase di progettazione, per questo motivo si stanno sviluppando sempre di più tecnologie innovative, come l'intelligenza artificiale (AI), che offrono supporto in tale ambito.

L'intelligenza artificiale rappresenta un supporto decisivo nell'adozione di strategie sostenibili, in quanto permettono di accelerare l'analisi di grandi quantità di dati, consentendo alle aziende di ottimizzare i propri processi.

Tra le tecnologie innovative che si stanno sviluppando, spiccano quelle basate sul Machine Learning (ML), sottoinsieme dell'intelligenza artificiale, le quali basano il loro funzionamento su metodi matematico-computazionali per apprendere informazioni direttamente da dati.

Lo scopo del presente lavoro è quello di proporre ed esporre la metodologia e gli strumenti utilizzati al fine di ottenere un modello riguardante l'impatto ambientale per componenti assialsimmetrici basato su tecniche di Machine Learning.

2 STATO DELL'ARTE

2.1 Sviluppo Sostenibile Nell'Industria 4.0

Al giorno d'oggi siamo di fronte alla Quarta Rivoluzione Industriale, detta anche Industria 4.0 (I4.0), dove tale termine venne coniato da tre ricercatori : Henning Kagermann, Wolf-Dieter Lukas e Wolfgang Wahlester in un convegno ritenuto a Hannover nel 2011. (Pfeiffer, 2017)

L'industria 4.0 mira a promuovere una crescente digitalizzazione delle imprese e dei loro sistemi produttivi, andando ad incentivare il ricambio generazionale dei processi, con l'obiettivo di riconvertirli in maniera "smart", questo grazie alle continue interconnessioni tra le tecnologie di produzione e di informazione.

Tuttavia, nonostante la grandiosa importanza di questo risultato, sembra essersi perso di vista quello che è, e dovrà essere un aspetto fondamentale come la sostenibilità.

Questo perché ancora oggi si è fatto ancora poco riguardante il tema della sostenibilità e l'industria 4.0 ha l'obbligo e il dovere di tener conto delle problematiche che la società moderna è afflitta come: la diminuzione e il progressivo esaurimento delle risorse, l'inquinamento ambientale, i cambiamenti climatici, ecc.

Il concetto di "sviluppo sostenibile" risale al 1987, dopo la pubblicazione del Rapporto di Brundtland (Gro, 1987), dove il concetto di sviluppo sostenibile è stato dato come, lo sviluppo che soddisfa i bisogni economici e sociali delle generazioni presenti senza andare a compromettere quelli delle generazioni future. Da tale definizione si evince però che il problema della sostenibilità è legato in maniera diretta alle necessità umane.

Nel ottica del Rapporto di Brundtland, la sostenibilità si basa su tre dimensioni:

- I. Sostenibilità ambientale, ovvero la capacità di conservare il capitale naturale, quindi in chiave industriale consiste nel mantenere la riproducibilità delle materie prime.
- II. Sostenibilità economica, ovvero la capacità di generare reddito e profitti
- III. Sostenibilità sociale, dove in chiave aziendale, si intende la capacità dell'impresa di curare con attenzione i rapporti con tutti i soggetti / individui / organizzazioni coinvolti nel progetto

[Digitare qui]

Industria 4.0 può rappresentare un ottimo aiuto rispetto allo sviluppo sostenibile, in particolare, l'obiettivo è quello di utilizzare il grande potenziale dell'Industria 4.0 per ottenere un modello sostenibile nelle sue tre dimensioni.

Per quanto riguarda la sostenibilità:

I. Ambientale: grazie all'Industria 4.0, si permette di ridurre i consumi di materie prime ed energia visto che se ne diminuisce l'utilizzo, d'altro canto la continua connessione e la comunicazione tra i dispositivi comporta un consumo di energia.

Grazie all'Industria 4.0 è possibile andare a lavorare sulla richiesta operativa in modo da andare a produrre solo lo stretto necessario e ridurre quindi gli scarti.

La riduzione degli scarti è poi collegata anche ai processi di *additive manufacturing*, ovvero la modalità di produzione che consente nella realizzazione di componenti attraverso principalmente stampanti 3D, quindi macchinari capaci di utilizzare solo la materia prima strettamente necessaria.

Un ulteriore fattore importante, utilizzabile grazie all'industria 4.0 è quello di poter memorizzare i dati relativi al ciclo vita di un prodotto in modo da facilitare l'eventuale riutilizzo produttivo.

II. Economica: grazie all'Industria 4.0, tramite l'automazione e la contemporanea operatività, permette di ridurre i tempi morti; inoltre è possibile ridurre gli errori e i difetti lungo la catena di montaggio andando così a ridurre costi dovuti proprio ad errori o scarti.

Grazie all'*additive manufacturing* e all'automazione flessibile, si permette la personalizzazione di componenti e la riduzione dei scarti e quindi l'aumento dei ricavi.

La produzione può essere organizzata:

- in lotti, dove la produzione è riservata a un numero ridotto di tipologie
- in serie, consiste nel produrre tipologie indipendenti tra loro
- Just In Time, consiste nel produrre la quantità giusta al momento giusto; ciò comporta la riduzione di sprechi

Per quanto riguarda quest'ultima tipologia di produzione (**JIT**), secondo (Erik Hofmann, 2017), è possibile tramite l'Industria 4.0 andare ad migliorare tale processo, ovvero dividendo il Just In Time in 4 fasi :

1. Production Planning
2. Production Order

[Digitare qui]

3. Disposition e Production
4. Delivery

III. Sociale: grazie all'Industria 4.0 si sono migliorate le condizioni di lavoro in termini di sicurezza, riducendo infortuni.

L'utilizzo dell'Industria 4.0, comporta numerosi benefici in termini di efficienza produttiva. Nonostante ciò, è necessario anche tener conto dei limiti e degli aspetti negativi che può comportare, come, la necessaria presenza di infrastrutture che assicurino una rete internet a banda larga, la tutela dei dati personali e le normative di tutela della privacy. Un altro limite, consiste nella carenza di risorse umane competenti. Un aspetto negativo è quello della completa automazione, che è vero che ha il vantaggio di ridurre energie e costi di produzione, ma ha lo svantaggio di ridurre posti di lavoro.

2.2 Life Cycle Assessment

L'analisi del ciclo di vita (Life Cycle Assessment – LCA) è una metodologia che permette di valutare gli impatti ambientali associati ad un prodotto, processo, o attività. L' LCA prende in esame l'intero ciclo di vita di un prodotto “dalla culla alla tomba” e il potenziale impatto ambientale in termini di uso e consumo delle risorse naturali (trattamento e lavorazione delle materie prime) e di conseguenze ambientali dei rilasci dovuti alla produzione, al trasporto, all'uso e alla manutenzione, fino al riciclo e alla collocazione finale.

L'analisi LCA è un mezzo che esprime in termini numerici il potenziale di impatto ambientale sui diversi macrosistemi terrestri attraverso l'identificazione e la quantificazione dei consumi di materia, energia e delle emissioni nell'ambiente e permette di confrontare in modo quantitativo prodotti simili e di individuare e valutare le opportunità per diminuire gli impatti.

L'analisi LCA può essere effettuata in qualsiasi fase del ciclo vita, ma la sua utilità è predominante nella fase iniziale di progettazione.

Le normative di riferimento sono: ISO 14040 del 2006 (ISO, 1404:2006 - Environmental management - LCA - Principles and Framework, 2006) , che contiene informazioni su concetti e procedure dell' LCA e la normativa ISO 14044 dello stesso anno (ISO, Environmental management-

[Digitare qui]

LCA- Requirements and Guidelines, 2006), che definisce nel dettaglio la procedura da seguire per svolgere un'analisi LCA.

LCA può aiutare a :

- Identificare opportunità atte a migliorare le prestazioni ambientali dei prodotti in vari punti del loro ciclo di vita
- Informare il progettista riguardo alle decisioni da prendere (ad esempio ai fini della pianificazione strategica, definizione delle proprietà, progettazione o riprogettazione di prodotti o processi)
- Selezionare gli indicatori pertinenti delle prestazioni ambientali, comprese le tecniche di misurazione, o nel marketing (ad esempio, implementare uno schema di etichettatura ecologica, presentare una dichiarazione ambientale o produrre una dichiarazione ambientale del prodotto)
- L'elenco seguente riassume alcune delle caratteristiche fondamentali della metodologia di valutazione del ciclo di vita:
- LCA esamina in modo sistematico gli aspetti ambientali e gli impatti dei sistemi di prodotto, dall'acquisizione delle materie prime allo smaltimento finale, in conformità all'obiettivo e al campo di applicazione definiti
- I risultati che si ottengono dal LCA sono legati all'unità funzionale definita, dal obiettivo e dai confini del sistema impostati
- Il grado di dettaglio e l'arco temporale del LCA possono variare ampiamente, in funzione della definizione dell'obiettivo e dai confini del sistema impostati
- La metodologia del LCA è aperta ad accogliere le nuove conoscenze scientifiche e lo stato dell'arte della tecnologia
- Requisiti specifici sono applicati al LCA, destinati ad essere usati per effettuare asserzioni comparative divulgate dal pubblico;
- Non esiste un unico metodo per condurre LCA
- LCA tratta i potenziali impatti ambientali; LCA non prevede impatti ambientali assoluti o precisi a causa di:
 - L'espressione relativa dei potenziali impatti ambientali a un'unità di riferimento
 - L'integrazione dei dati ambientali nello spazio e nel tempo;
 - L'incertezza intrinseca nella modellazione degli impatti ambientali
 - Il fatto che alcuni possibili impatti ambientali sono chiaramente impatti futuri

[Digitare qui]

- La fase di valutazione dell'impatto (Life Cycle Impact Assessment -- LCIA), unitamente ad altre fasi LCA, fornisce una prospettiva a livello di sistema dei problemi ambientali e delle risorse per uno o più sistemi di prodotto
- LCIA assegna i risultati della fase di analisi di inventario (Life Cycle Inventory –LCI) alle categorie d'impatto; per ogni categoria è selezionato un indicatore di categoria di impatto del ciclo di vita ed è calcolato il risultato di tale indicatore (risultato dell'indicatore); la raccolta dei risultati degli indicatori (risultati di LCIA) o il profilo LCIA fornisce le informazioni sui problemi ambientali associati agli elementi di ingresso e in uscita del sistema prodotto
- Non esistono basi scientifiche per la riduzione dei risultati dell'LCA a un singolo punteggio o numero complessivo, poiché la ponderazione richiede la scelta di valori
- L'interpretazione del ciclo di vita richiede l'uso di un procedimento sistematico per identificare, qualificare, verificare, valutare, o presentare le conclusioni sulla base dei risultati del LCA, allo scopo di soddisfare i requisiti descritti nell'obiettivo e nel campo di applicazioni definiti nella fase iniziale dello studio
- L'interpretazione del ciclo di vita richiede l'uso di un procedimento iterativo sia per la fase di interpretazione e sia per le altre fasi dell'LCA
- L'interpretazione del ciclo di vita prevede l'individuazione dei collegamenti tra LCA e altre tecniche di gestione ambientale, sottolineando i punti di forza e i limiti del LCA in relazione alla definizione del suo obiettivo e campo di applicazione
- La metodologia del LCA definisce le quattro fasi di seguito specificate

1. **DEFINIZIONE DELL'OBIETTIVO E DEL CAMPO DI APPLICAZIONE:** si definisce, in primo luogo, lo scopo dell'analisi (come ad esempio il carico ambientale in un determinato ecosistema, valutazione delle emissioni di CO₂ ecc.); poi si definisce l'unità funzionale, ovvero l'azione che sarà soggetta all'analisi. La funzione è dotata di unità di misura ed è il parametro di riferimento su cui si basa tutta l'analisi. In questa fase poi definiscono le categorie di impatto ambientale, le metodologie di calcolo degli impatti, qualità dei dati e assunzioni fatte.

Come ultimo definisco i confini del sistema, ovvero le fasi che verranno incluse nell'analisi

I confini del sistema possono comprendere:

- Fase Material: estrazione e processazione dei materiali
- Fase Produzione: lavorazione dei materiali per produrre i componenti e il loro assemblaggio

[Digitare qui]

- Fase d'Uso: utilizzo del prodotto che permette di valutare oltre l'eventuale manutenzione, anche l'impatto ambientale derivante dal consumo energetico richiesto
 - Fase Fine Vita: questa fase comprende gli scenari che si intendono attuare per il corretto smaltimento del prodotto, che riguardano sia gli open-loop scenario (incenerimento o deposito discarica) sia i closed-loop scenario (riutilizzo rigenerazione e riciclo) ; le operazioni di disassemblaggio vengono introdotte in questa fase
2. LCI (Life Cycle Inventory): costituisce la fase di reperimenti di tutti i dati necessari per lo svolgimento dell'analisi a seconda dell'obiettivo precedentemente definito
Questa fase è la fase che richiede più attenzione perché ne dipende l'accuratezza dell'analisi stessa, possiamo identificare alcuni passaggi principali
- Considerazione dell'obiettivo
 - Preparazione della raccolta dati
 - Collezionare i dati (dall'azienda ,dai clienti, dai fornitori, da letteratura)
 - Validare i dati
 - Relazionare i dati all'unità funzionale
 - Raffinare, se necessario (ad esempio per mancanza di dati)i confini del sistema
 - Revisionare e ripetere se necessario alcuni passaggi
3. VALUTAZIONE DEGLI IMPATTI :si determinano gli impatti ambientali dell'oggetto di valutazione. Questa fase è solitamente accompagnata da software commerciali. I risultati devono essere presentati in maniera pertinente e comprensibile
4. **INTERPRETAZIONE DEI RISULTATI:** si analizzano i risultati finali anche attraverso confronti con risultati presenti in letteratura, in modo tale da redigere raccomandazioni indirizzate al progettista o in generale perseguire l'obiettivo per cui è stata svolta l'analisi.

Si sottolinea che il compito dell'analisi LCA non è dare soluzione al problema, il quale viene lasciato alla capacità del progettista, ma ha come obiettivo quello di analizzare

[Digitare qui]

le criticità di carattere ambientale del ciclo vita del prodotto e, alla luce dei risultati, proporre dei consigli.

2.3 Machine Learning Per La Sostenibilità Ambientale

Nell'era dell'Industria 4.0, la produzione è strettamente interconnessa con i sistemi di informazione e comunicazione, rendendola più competitiva e con maggiore versatilità.

Tecnologie come l'Intelligenza Artificiale (AI), Internet of Things (IoT), Big Data (BD), e Machine Learning (ML), forniscono prospettive fondamentali per la crescita e l'innovazione aziendale.

Il Machine Learning è un campo di studio che esamina l'utilizzo di algoritmi computazionali per trasformare dati empirici in modelli utilizzabili. Gli algoritmi di Machine Learning, possono essere utilizzati per:

- a. Raccogliere la comprensione del fenomeno che ha prodotto i dati in studio
- b. Estrarre la comprensione dei fenomeni sotto forma di un modello
- c. Prevedere i valori futuri di un fenomeno, utilizzando il modello precedentemente ottenuto
- d. Rilevare un comportamento anomalo esibito da un fenomeno osservato

Risulta quindi possibile utilizzare il Machine Learning e l'elaborazione dei dati in tempo reale per migliorare l'efficienza, semplificare i processi produttivi e promuovere lo sviluppo.

La produzione entrerà pienamente in un'era sostenibile con l'ausilio dell'Industria 4.0

[Digitare qui]

2.4 Casi Studio

In questi paragrafi , si mostreranno alcuni esempi di applicazione del Machine Learning a riguardo della sostenibilità ambientale.

2.4.1 Costruzione Di Un Sistema Di Trasporto Intelligente a Basse Emissioni Di Carbonio Basato Su BD e ML

Secondo ricerche, le emissioni di gas di scarico dei veicoli sono una parte importante di inquinamento ambientale ,inoltre, il grande aumento del numero di automobili, ha peggiorato questa situazione.

Nella trattazione (Tu Peng, 2020), viene proposto un algoritmo su base Big Data e Machine Learning , che considera il consumo del carburante e le emissioni di scarico dei veicoli.

Di solito, i metodi della previsione del flusso del traffico sono basati su metodi matematici tradizionali, ovvero grazie all'esistenza di una legge di andamento del flusso di traffico, ne permette la previsione. Tuttavia tali metodi non permettono di prevedere cambiamenti in tempo reale del flusso, soprattutto in larga scala con un elevata precisione.

Per questo motivo, molti ricercatori si sono rivolti a metodi di Machine Learning o metodi di rete neurale, con lo scopo di rappresentare in modo più preciso il flusso di traffico in tempo reale.

Al fine di ottenere tale risultato, bisogna anche tener conto anche del algoritmo che considera il percorso ottimale, l'obiettivo di tale algoritmo è quello di pianificare il percorso ottimale, ovvero trovare il percorso ottimale dal punto di partenza a quello di arrivo, con la restrizione di minimizzare la funzione costo. A tal proposito, (McGinty & Smyth, 2000) , ne esegue uno studio.

L'emissione di gas di scarico di un veicolo è correlata a molti fattori, come la situazione specifica di quel determinato veicolo, la situazione specifica della rete stradale, il traffico della rete stradale, la velocità di marcia del veicolo, la distanza percorsa e così via; visto che maggiore sarà la distanza del viaggio e maggiore sarà il consumo, oppure maggior traffico anche comporterebbe maggiori consumi. Questo fa sì che algoritmo è costituito da tre modelli principali:

1. Un modello semplificato riguardante il consumo del veicolo
2. Un modello di previsione del flusso di traffico

[Digitare qui]

3. Un modello di navigazione a basse emissioni di carbonio

Utilizzando l'algoritmo ottenuto, è stato dimostrato come percorrendo 2,6 km, si è ottenuto un minor consumo di carburante e una minore emissione di gas di scarico rispetto ad altri algoritmi. Tuttavia l'algoritmo ottenuto percorre un percorso più sfavorito in termini di tempo, ovvero 6 secondi in più rispetto a quello ottimale.

Questo risultato per percorsi molto più lunghi potrebbe essere quindi amplificato.

Questo studio ha dimostrato come poter aiutare a ridurre le emissioni dei veicoli, ma con l'aspetto negativo in termini di tempo; in futuro comunque si condurranno ulteriori ricerche per ottenere una soluzione migliorata.

2.4.2 Approccio Basato Sui Dati Per La Scoperta Dei Potenziali Di Risparmio Energetico Nella Fabbrica Di Produzione

Negli ultimi anni, ci si sta dirigendo sempre di più verso lo sviluppo sostenibile, come si può dimostrare dall'accordo di Parigi del 2015, (krivosheeva OI, 2022), dove i paesi che ne hanno preso iniziativa, hanno iniziato a modificare le disposizioni legislative e introdurre nuove normative atte a ridurre le emissioni di gas a effetto serra (GHG) .

Visto che l'industria risulta essere uno dei principali consumatori di combustibili fossili, è importante andare a cercare di migliorare il più possibile l'efficienza energetica.

In questo studio, (Bin Song, 2018), è stato proposto un approccio di analisi dei dati per andare a migliorare l'efficienza energetica ed andare a quantificare i potenziali risparmi.

In generale, una fabbrica di produzione presenta due tipi di attrezzature:

- Attrezzature per la produzione, come macchine utensili
- Attrezzature ausiliare, come refrigeratori, caldaie, illuminazioni, ecc

Dallo studio (Katili AR, 2015), si osserva che tra tutta l'energia consumata, circa 55% è dovuto all'aria condizionata, circa il 25% da compressori d'aria, il 15% circa da stampaggio ad iniezione di materiale plastico e il restante dall'illuminazione; quindi più dei 2/3 del consumo energetico è dovuto ad attrezzature ausiliare.

[Digitare qui]

Occorre, eseguire un approccio di analisi dei dati per scovare inefficienze energetiche, per fare ciò si fa riferimento al concetto di consumo di potenza come parametro da confrontare per le diverse macchine, le quali avranno diversi input e output , che spesso sono diversi tra loro. Tali informazioni verranno inserite all' interno di una matrice insieme a serie temporali.

La matrice esegue un analisi di Pearson, che verrà trattata in seguito, dove si determina la correlazione tra i parametri necessari con l'output.

Nei programmi di miglioramento dell'efficienza energetica, uno dei obiettivi è migliorare le unità organizzative come reparti, gruppi, ecc . Ciascuna unità organizzativa ha la sua linea di base e le sue potenzialità dovute a differenze funzionali e sforzo passato, avere una linea di base troppo bassa fa sì, che si stimola poco sforzo, troppo alta comporta a demotivare il personale. Viene definita frontiera la linea di base che fa sì che si ottiene la massima efficienza energetica.

Le pratiche comuni in tali settori consistono nella sostituzione con una macchina più efficiente dal punto di vista energetico o l'applicazione di nuove tecnologie se economicamente possibile.

Tuttavia è possibile andare a utilizzare un opportuno algoritmo di Machine Learning per ridurre al minimo l'inefficienza.

Per ricapitolare, quello che è stato fatto in questo studio basato sull'identificazione di opportunità di risparmio energetico e sulla quantificazione dell'eventuale risparmio è stata:

- 1) Preparazione dei dati
- 2) Analisi di correlazione (tramite Pearson)
- 3) Analisi dei dati
- 4) Quantificazione del potenziale di risparmio energetico

Per dimostrarne l'utilità , tale sistema è stato applicato ad un sistema di refrigerazione, dove si è ottenuto che l'output del refrigerante è correlato a :

- Temperatura dell'acqua refrigerante in ingresso (°C)
- Temperatura dell'acqua al ingresso del condensatore (°C)
- Consumo energetico (kW)

Il Risultato ottenuto comporta un efficienza energetica del 8,6%.

[Digitare qui]

3 METODO

In questo capitolo si illustreranno le fasi fondamentali e gli strumenti necessari per determinare un modello di impatto ambientale per un componente meccanico.

3.1 Tecniche Di Stima

Le tecniche di stima utilizzate in un progetto possono essere utilizzate per determinare diversi parametri.

Gli approcci e le strategie di stima possono essere ricondotti a due categorie

- Metodo di stima analitica
- Metodo di stima strategica

Le due tecniche di stima possono essere applicate in alternativa (in funzione delle informazioni disponibili) oppure in combinazione in modo da arrivare a valutazioni congrue.

Indipendentemente dal tipo di approccio seguito, possono essere applicate:

- Tecnica di stima analogica
- Tecnica di stima parametrica
- Tecnica di stima dei tre punti
- Tecnica di stima intuitiva

3.1.1 Tecniche Di Stima Intuitiva

Le tecniche intuitive di stima si basano sull'esperienza passata di un esperto o un team di esperti.

La stima prevede la formulazione di un parere da parte di soggetti componenti nell'ambito che si rende necessario affrontare, nel nostro caso l'impatto ambientale. Può essere svolta attraverso un coinvolgimento diretto oppure indiretto applicando la tecnica Delphi.

[Digitare qui]

Tale tecnica risulta molto utile quando:

- Si ha bisogno di fare previsioni a lungo termine oppure quando si deve affrontare una nuova situazione dove la opinione degli esperti è l'unica fonte di informazione disponibile
- Necessità di far emergere nuove opzioni/alternative
- Necessità di un parere indipendente o esterno per prendere alcune decisioni.

Per applicare la tecnica Delphi, si rendono necessarie alcune condizioni:

- Procedere per cicli di approfondimento successivo in modo da isolare le questioni più rilevanti e svilupparle in profondità
- Garantire l'indipendenza di giudizio attraverso l'anonimato di ciascun esperto in modo che ciascuno sia libero di esprimere il proprio pensiero

Queste conoscenze possono inoltre essere archiviate in un luogo specifico, come per esempio un database, sotto forma di regole o giudizi in modo da rendere più facili e standardizzate le stime di componenti futuri.

3.1.2 Tecnica Di Stima Analogica

Le tecniche di stima analogiche consistono nell'utilizzo di dati storici di progetti già completati ed eseguendo un confronto con il progetto in esame, cercando di identificare delle similarità. I criteri di somiglianza più utilizzati sono due:

1. Modelli di analisi di regressione; in tali modelli i dati storici riguardanti l'impatto ambientale sono utilizzati con lo scopo di stabilire una relazione lineare tra i dati riguardanti l'emissione del passato e il valore di qualche variabile degli stessi. Le variabili vengono scelte in maniera tale che possono essere utili a prevedere le emissioni future.
2. Modelli di rete neurale di back-propagation (BPNN). In tali modelli si utilizza una rete neurale (NN), la quale se correttamente addestrata, è in grado di archiviare la conoscenza per dedurre risposte a domande che non erano state poste alla rete durante il suo addestramento

[Digitare qui]

3.1.3 Tecnica Di Stima Analitica

Questa metodologia di tecnica basa il suo funzionamento sullo scomporre il prodotto in unità elementari, attività ed operazioni, le quali risulteranno coincidere con le fasi del ciclo produttivo a cui è sottoposto il prodotto preso in esame e di stimarne le risorse desiderate; le risorse totali saranno la somma di tutte le risorse.

Tale tecnica risulta essere quella utilizzata dal software SimaPro, che è quello utilizzato in questo studio

3.1.4 Tecnica Di Stima Parametrica

Le tecniche di stima parametriche risultano essere tra le più accurate; risultano inoltre essere particolarmente efficaci quando sono ben definiti i parametri che influiscono maggiormente l'impatto ambientale.

Come già detto, la tecnica di stima parametrica, rispetto alla stima analogica, risulta però utilizzabile solo dopo aver individuato alcuni parametri e ideato un algoritmo necessario nel compiere i calcoli.

La differenza rispetto alla stima analitica è che la stima parametrica è vero che ha bisogno anche lei di dati storici, ma tali dati storici non devono essere necessariamente provenienti da prodotti molto simili rispetto a quelli alla quale si vuole valutare l'impatto ambientale.

3.2 Metodo Per Lo Sviluppo Di Un Modello D' Impatto Ambientale

Gli strumenti utilizzati per lo sviluppo di un modello di impatto ambientale sono:

- SimaPro. il quale esegue una stima di tipo analitico
- RapidMiner, il quale esegue un'analisi parametrica

Per quanto riguarda il primo strumento, ovvero SimaPro il quale basa il suo funzionamento mediante tecniche analitiche ed è stato utilizzato per definire stime di impatto ambientale sia per le diverse tipologie di lavorazioni, sia per i diversi materiali studiati.

[Digitare qui]

Ricordando che le metodologie analitiche, sono in grado di scomporre il prodotto in un certo numero di fasi e operazioni elementari che vanno a costituire il ciclo produttivo del prodotto stesso.

Queste operazioni sono fondamentali per il corretto funzionamento del secondo strumento (RapidMiner), in quanto hanno permesso di fornire lo storico riguardante i dati dei diversi componenti.

Le informazioni ottenute tramite SimaPro verranno usate per creare un database formato da un elevato numero di “records”, il quale farà sì che si otterrà un ottimo modello di impatto ambientale .

3.3 Strumenti Utilizzati

3.3.1 SimaPro

SimaPro è la soluzione software LCA leader, con una reputazione di 30 anni nell’industria e nel mondo accademico in più di 80 paesi.

Questo software è stato progettato per essere una fonte di informazioni scientifiche, fornendo piena trasparenza ed evitando processi black-box. Esso inoltre è in grado di far prendere decisioni consapevoli durante l’analisi, in modo da garantire l’accuratezza dei risultati.

SimaPro è un software per la valutazione del ciclo di vita (LCA) in accordo con le norme ISO; esso permette di raccogliere, analizzare e monitorare i dati relativi alle prestazioni ambientali di prodotti e servizi.

Tramite SimaPro è possibile:

- Modellare e analizzare facilmente cicli di vita complessi in modo sistematico e trasparente
- Misurare l’impatto ambientale dei tuoi prodotti e servizi in tutte le fasi del ciclo di vita
- Identificare i punti critici in ogni fase di produzione, dall’estrazione delle materie prime alla produzione, utilizzo e smaltimento

SimaPro dispone di un database completo e personalizzato di processi, materie prime e macchine, il quale permette l’automazione di numerose attività

[Digitare qui]

3.3.2 RapidMiner

RapidMiner, precedentemente noto come YALE (Yet Another Learning Environment), è una piattaforma software generica per la preparazione dei dati, machine learning, text mining, e l'analisi predittiva, utilizzando un metodo client/server. (Patrick Schneider, 2022)

Il metodo client/server consiste nel suddividere l'attività lavorative tra i fornitori di una risorsa o di un servizio, detti server e quelli che richiedono il servizio, ovvero i client.

Di solito risulta che client e server non comunicano su una rete computer ma su hardware separato oppure sia il client che il server possono risiedere nello stesso sistema.

RapidMiner fornisce procedure di data mining e machine learning tra cui:

- Caricamento dei dati
- Trasformazione dei dati
- Pre elaborazione e visualizzazione dei dati
- Analisi predittiva
- Modellazione statistica
- Valutazione
- Distribuzione

Le caratteristiche di RapidMiner si possono riassumere nel fatto che, tramite un'interfaccia user friendly (facilmente usabile), permette di progettare il processo di analisi dei dati, i quali possono essere sia numerici che testuali, e si potranno poi importarli per rielaborarli ed analizzarli meglio, anche creando altri file di output con il risultato dell'analisi.

Una volta effettuata l'analisi di impatto ambientale è possibile graficare l'analisi effettuata tramite due grafici:

- Nel grafico superiore, quello di *Figura 3.1* è possibile valutare l'influenza di ciascun parametro con l'output
- Nel grafico inferiore, ovvero quello di *Figura 3.2*, invece è possibile valutare quanto si discosta il valore effettivo rispetto al valore predetto

[Digitare qui]

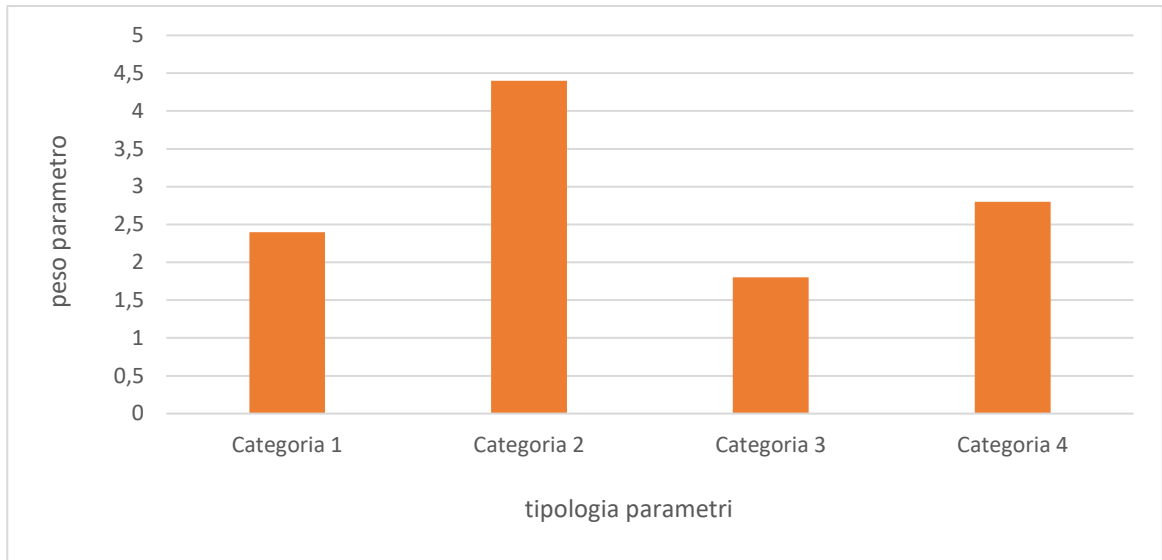


Figura 3 . 1 Peso parametro rispetto all'output

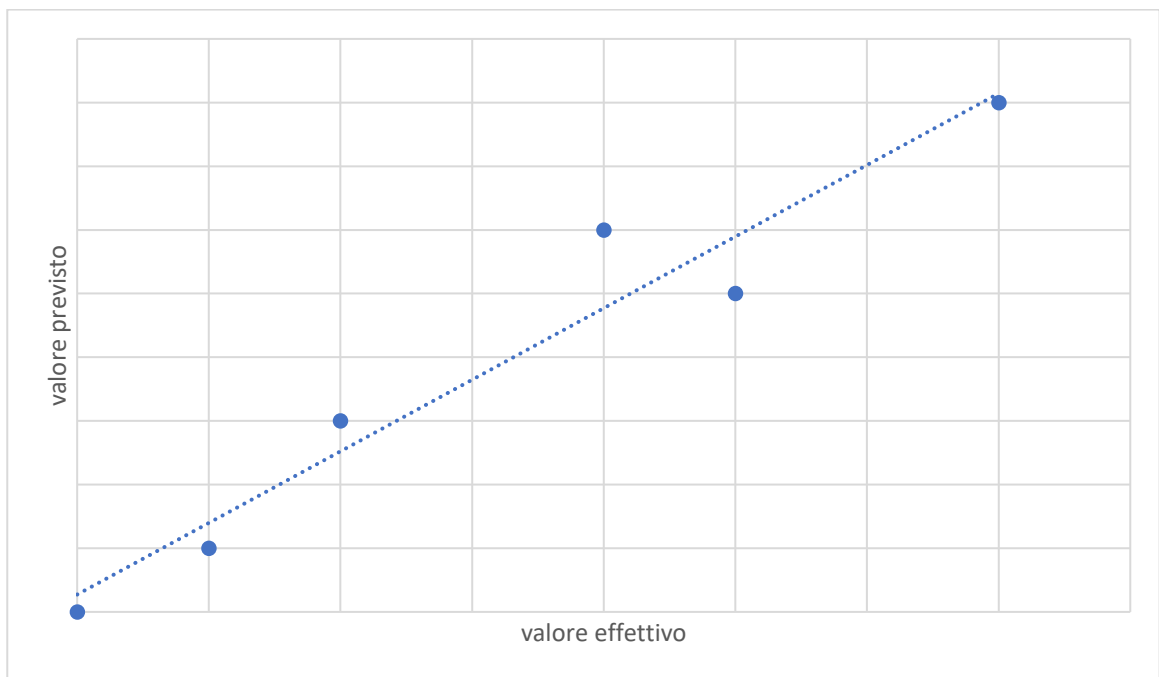


Figura 3 . 2 Prediction Chart

[Digitare qui]

3.4 Creazione Database

Sapendo che lo strumento di stima parametrica RapidMiner ha bisogno di un database, esso sarà di particolare importanza, visto che il modello di stima parametrica ha bisogno di dati storici riguardanti i componenti.

Quindi possono essere utilizzati dei componenti di "riferimento" e da questi componenti è possibile estrapolare e registrare le informazioni, ovvero parametri non geometrici e geometrici contenenti anche le informazioni riguardanti il loro impatto ambientale.

Dunque un generico database è strutturato come in *Tabella 1* , ovvero è formato da :

- N°VARIABILI INDIPENDENTI: numero di parametri utilizzati per definire il modello di impatto ambientale. In questo caso sono i parametri geometrici e non geometrici, e corrispondono alle colonne del mio database
- N° di "RECORDS" : numero di elementi da cui ricavare le variabili indipendenti. In questo caso sono il numero di componenti, e corrispondono al numero di righe.

Tabella 1 Esempio Database

	A	B	C
1			
2			
3			

N° VARIABILI INDIPENDENTI

N° RECORDS

[Digitare qui]

Nel caso studiato, il numero di variabili indipendenti costituisce il numero di parametri geometrici e non dei vari componenti, mentre il numero di records rappresenta il numero di componenti presi in esame.

Andranno dunque riportati su ciascuna riga di un foglio Excel, tutti i componenti esaminati nel nostro studio, mentre su ciascuna colonna tutti i parametri dei componenti, ricavati tramite SimaPro, ritenuti influenti rispetto al parametro da predire, ovvero l'impatto ambientale.

I parametri da selezionare per effettuare una buona stima parametrica dell'impatto ambientale, devono per prima cosa influenzare quest'ultimo e poi devono essere indipendenti, ovvero avere una bassa correlazione tra di loro.

Per rendere più efficiente l'analisi, si può eseguire l'analisi di Pearson, la quale verrà successivamente trattata, che ha lo scopo di andare ridurre il numero di parametri fino ad ottenere le variabili indipendenti che influiscono maggiormente l'impatto ambientale.

3.5 Analisi Di Pearson

L'analisi di Pearson permette di andare a ridurre il numero di variabili utilizzati come parametri per andare ad individuare il modello di impatto ambientale, in modo che rimangono solo variabili strettamente legate alla variabile dipendente in esame, ovvero l'impatto ambientale, e che queste siano indipendenti.

Per poter eseguire questa analisi, bisogna prima verificare che le variabili superino una checklist composta da 5 controlli, solo se le variabili superano questi 5 controlli, allora è possibile utilizzare la correlazione di Pearson per ottenere dei risultati validi per la nostra analisi.

I controlli sono i seguenti:

1. LE DUE VARIABILI DEVONO ESSERE ENTRAMBE QUANTITATIVE: una variabile si definisce quantitativa quando le sue determinazioni (valori) sono espresse mediante valori numerici
2. LE DUE VARIABILI QUANTITATIVE DEVONO ESSERE ABBINATE SUGLI STESSI CASI : questo vuol dire che per ogni componente deve essere misurato un valore sia per una variabile che per l'altra

[Digitare qui]

3. IL GRAFICO DI CORRELAZIONE DEVE MOSTRARE UNA RELAZIONE DI TIPO LINEARE TRA LE DUE VARIABILI: come si può notare dalla *Figura 3.3*

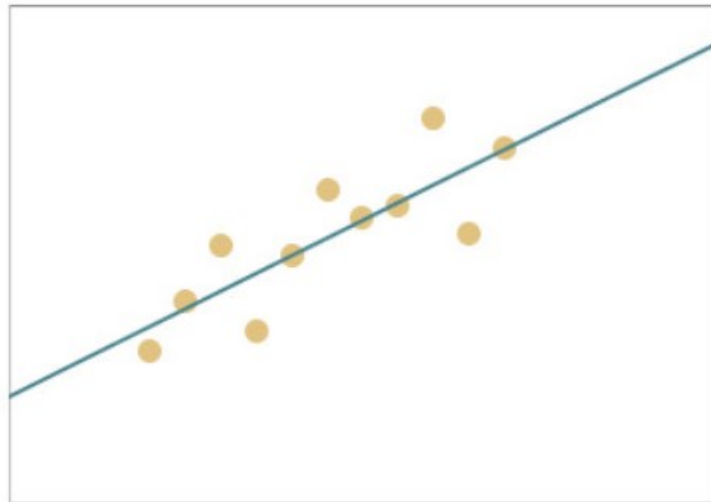


Figura 3.3 Grafico di correlazione lineare

- Qualora non avesse una relazione di tipo lineare, bisogna verificare che sia almeno monotona, ovvero se all'aumentare dei valori di una variabile, i valori dell'altra variabile aumentano o diminuiscono anche in modo non lineare; se questo è verificato è possibile valutare la relazione tramite l'indice di correlazione di Spearman (Paola, 2020)
4. NON DEVONO ESSERE PRESENTI OUTLIERS INFLUENTI: gli outliers sono dei casi all'interno del campione che hanno caratteristiche molto diverse rispetto alle altre osservazioni
 5. LA DISTRIBUZIONE DI ENTRAMBE LE VARIABILI DEVE ESSERE NORMALE: tale verifica può essere fatta attraverso un test di normalità

L'analisi di Pearson basa il suo funzionamento sul concetto di correlazione tra variabili (Pozzolo, 2020), questa correlazione è basata tramite un indice, detto indice di correlazione ("r").

L'indice "r" sarà sempre compreso tra -1 (che indica la perfetta relazione negativa lineare tra le due variabili) e +1 (che indica la perfetta relazione positiva lineare tra le due variabili), come si può vedere in *Figura 3.3*.

[Digitare qui]

Invece un valore di indice pari a 0, corrisponde al fatto che non è presente nessuna relazione tra le due variabili.

Nella realtà il valore di indice pari esattamente a -1 e +1 non potrà mai capitare in quanto è un limite puramente teorico.



Figura 3 . 4 Correlazione dell'analisi di Pearson

Per definizione risulta che due variabili quantitative sono correlate :

- Positivamente, se quando le unità statistiche che hanno alti/bassi valori di una variabile, tendono ad avere alti/bassi valori per l'altra variabile
- Negativamente, se quando le unità statistiche che hanno alti/bassi valori di una variabile, tendono ad avere bassi/alti valori per l'altra variabile

Il primo passo che effettua l'analisi di Pearson per ridurre il numero di parametri è quello di valutare il grado di correlazione tra tali parametri e l'impatto ambientale, tramite l'indice di correlazione "r". Una volta ottenuto l'indice di correlazione per tutti i parametri, si eliminano le variabili che hanno una bassa correlazione con l'impatto ambientale, sono considerate situazioni di bassa correlazione quando si riscontra un indice $r < 0,3$.

Dopo di ciò, si valuta il grado di correlazione tra un parametro e l'altro, in modo da far sì che rimangano solo i parametri indipendenti, ovvero quelli che hanno una bassa correlazione tra di loro, e per fare ciò si eliminano i parametri che hanno un indice di correlazione pari a $r > 0,8$.

[Digitare qui]

3.6 Analisi Machine Learning

Lo scopo dell'apprendimento è fondamentale per capire quali tecniche bisogna utilizzare, bisogna capire la tipologia di ciò che vogliamo predire per applicare i corretti algoritmi ed avere quindi un risultato corretto.

Per fare ciò, bisogna porsi la domanda: « di che tipo è il valore di ciò che voglio predire? », in base alla risposta a tale domanda, quindi in base al valore che vogliamo predire, si hanno a disposizione diverse aree di analisi.

Se quello che vogliamo predire è una variabile di tipo:

- Categorico, cioè può assumere solo una serie finita di valori, allora l'area di analisi è la classificazione

- Numerico continuo, ovvero un numero che può assumere infiniti valori e l'area di analisi è quella di regressione

Infine risulta esserci un'ulteriore divisione, è un raggruppamento e non una classificazione, utilizzabile quando lo scopo dell'analisi è quello di dividere le osservazioni in gruppi (clustering). La differenza è che nelle classificazioni si classificano in base a categorie che sappiamo inizialmente, mentre nel clustering non sei a conoscenza.

Una volta ottenuto il database definitivo, è necessario suddividere per RapidMiner i dati in set di "training" e set di "testing".

I dati utilizzati nella fase di "testing" sono quelli che su cui viene eseguita la stima, mentre quelli utilizzati nella fase di "training" servono ad "allenare" l'algoritmo che fornisce la stima.

Una percentuale tipica di suddivisione di tali dati è 80% per la fase di training e il restante, ovvero il 20% nella fase di "testing"

Le stime dell'impatto ambientale predette da RapidMiner, vengono poi confrontate con i valori effettivi tramite l'errore di stima.

[Digitare qui]

Questo errore viene misurato tramite il MAPE (Mean Absolute Percentage of Error) :

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

Dove:

y_i è il *valore effettivo*

\hat{y}_i è il *valore predetto*

Ad un basso valore di MAPE corrisponde una miglior stima parametrica

Il MAPE è la media aritmetica dei rapporti tra il valore assoluto degli errori di previsione e la domanda effettiva verificatasi in n intervalli (tutti di uguale durata) fino ad arrivare al periodo rispetto quale è stata effettuata la previsione.

Il MAPE mette in relazione l'errore con il valore della domanda, in questo modo è possibile "mediare" l'errore con la sua rilevanza effettiva.

Un vantaggio del MAPE è che può essere utilizzato come una buona misura di accuratezza poiché è indipendente dalla scala e può essere utilizzato per confrontare diverse serie o scenari previsionali.

Per quanto riguarda sempre il MAPE, risulta avere lo svantaggio che diventa indefinito quando il valore effettivo è nullo per qualsiasi osservazione dei dati

[Digitare qui]

4 CASO STUDIO

In questo capitolo si esporrà l'applicazione della metodologia precedentemente descritta, andando a vedere nel dettaglio quali sono componenti presi in esame nel caso studio e come è stato possibile, nel dettaglio, effettuare tutti gli step precedentemente descritti fino ad ottenere il modello di impatto ambientale.

4.1 Componenti Del Caso Studio

I componenti presi in esame per lo sviluppo di un modello di impatto ambientale sono 73, i quali risultano essere:

- Componenti meccanici (alberi, dischi ecc.)
- Assialsimmetrici , non solo per quanto riguarda la geometria, ma anche: carichi, vincoli, proprietà del materiale, sono simmetrici rispetto un asse.

Per ogni componente, si andrà a valutare il suo impatto ambientale per tre diverse tipologie di materiale, ovvero si andrà a valutare nel dettaglio:

- Aluminium cast alloy
- Chromium steel 18/8
- Steel Low-Alloyed

Quindi si avranno a disposizione 219 modelli virtuali, più precisamente modelli CAD 3D.

Il CAD (*“Computer Aided Design”*) indica un settore dell'informatica volto all'utilizzo di tecnologie software e in particolare alla computer grafica per supportare l'attività di progettazione. I sistemi di *Computer Aided Design* hanno come obiettivo la creazione di modelli, soprattutto 3D, dei manufatti.

Per valutare l'impatto ambientale di tali componenti è necessario anche tener conto delle operazioni che essi subiscono, ed avendo come campioni solo pezzi assialsimmetrici, si andranno a tener conto solo delle operazioni di:

- Tornitura: è una lavorazione per asportazione di materiale , essa è definita da un moto rotatorio del pezzo e un moto per lo più rettilineo del utensile. Solitamente la lavorazione avviene in due passaggi: nel primo, chiamato sgrossatura, si rimuove il “grosso” del sovrametallo

[Digitare qui]

tramite utensili in grado di resistere a forti sollecitazioni; nella seconda fase, chiamata finitura, il pezzo viene lavorato al fine di ottenere una superficie finale con dimensioni, tolleranze e rugosità previste.

- Fresatura: è una lavorazione per asportazione di materiale mediante l'utilizzo di un utensile tagliente a geometria definita. La fresatura richiede la rotazione dell'utensile e la traslazione del pezzo.

Il ciclo lavorativo prevede normalmente una prima fase di sgrossatura, in cui l'asportazione viene fatta nel modo più rapido e quindi più economico possibile, lasciando un sufficiente strato di sovrametallo per la successiva fase di finitura in cui si asportano gli ultimi strati di metallo in modo da ottenere una superficie finale con dimensioni, tolleranze e rugosità previste inizialmente da progetto

- Foratura: con tale operazione si intende la lavorazione che consiste nell'eseguire fori di diverso diametro e diversa profondità.

I fori vengono eseguiti tramite il moto rotatorio della punta elicoidale lungo il proprio asse e tramite la traslazione longitudinale dell'utensile lungo lo stesso asse.

[Digitare qui]

4.2 Analisi Del Impatto Ambientale Tramite SimaPro

Tramite SimaPro è possibile valutare l'impatto ambientale, ovvero l'emissione di CO₂ dei materiali (*Tabella 2*) e delle lavorazioni compiute (*Tabella 3*) per eseguire i modelli.

Di seguito vengono allegati i valori ricavati tramite tale analisi:

Tabella 2 Impatto Ambientale dei Materiali

MATERIALE	EMISSIONI DI CO₂ [kgCO₂ eq]
ALLUMINIUM CAST ALLOY	5,476950712
CHROMIUM STEEL 18/8	4,51614
STEEL LOW-ALLOYED (S235JR)	1,651793

Le emissioni di CO₂ sopra riportate sono riferite al kg di materiale prodotto.

Tabella 3 Impatto Ambientale delle Lavorazioni

MATERIALE	TORNITURA [kgCO₂ eq]	FRESATURA [kgCO₂ eq]	FORATURA [kgCO₂ eq]
ALLUMINIUM CAST ALLOY	12,73157225	12,89009272	12,54330728
CHROMIUM STEEL 18/8	6,907024387	7,1319129512	6,695018119
STEEL LOW- ALLOYED	3,897320655	4,143878128	3,787803865

Le emissioni di CO₂ sopra riportate sono riferite al kg di materiale asportato.

[Digitare qui]

4.3 Raccolta Dati

Per ogni singolo componente, vengono raccolti in un file Excel, le informazioni e i dati ritenuti utili per la successiva fase di analisi tramite RapidMiner.

Ovvero:

- Dati geometrici e non geometrici dei componenti, ottenuti dall'analisi dei modelli CAD, e sono:
 - Materiale
 - Massa
 - Tipologia di finito: se si tratta di un pieno o cavo
 - Dimensioni di ingombro D1,D2 e D3, dove le prime due dimensioni rappresentavano le massime estensioni lungo la direzione radiale, mentre l'ultima dimensione rappresenta l'estensione in direzione assiale.
 - Diametro interno (nullo in caso il componente non presenta cavità)
 - Volume
 - Superficie totale
 - Numero superfici
 - Numero smussi
 - Numero raccordi
 - Numero fori
 - Presenza di particolarità nella geometria del componente come:
 - Sedi per anelli
 - Sedi per linguetta
 - Sedi per chiavetta
 - Presenza di geometrie non coassiali

- Dati ottenuti dal analisi tramite SimaPro, ovvero:
 - Emissioni di CO2 eq per ogni kg di Alluminium Cast Alloy prodotto

[Digitare qui]

- Emissioni di CO₂ eq per ogni kg di Chromium Steel 18/8 prodotto, dove il numero 18 rappresenta la percentuale di Cromo e 8 indica la percentuale di Nichel
- Emissioni di CO₂ eq per ogni kg di Steel Low Alloyed
- Emissioni di CO₂ eq per ogni kg di Alluminium Cast Alloy asportato tramite tornitura
- Emissioni di CO₂ eq per ogni kg di Alluminium Cat Alloy asportato tramite fresatura
- Emissioni di CO₂ eq per ogni kg di Alluminium Cast Alloy asportato tramite foratura
- Emissioni di CO₂ eq per ogni kg di Chromium Steel 18/8 asportato tramite tornitura
- Emissioni di CO₂ eq per ogni kg di Chromium Steel 18/8 asportato tramite fresatura
- Emissioni di CO₂ eq per ogni kg di Chromium Steel 18/8 asportato tramite foratura
- Emissioni di CO₂ eq per ogni kg di Steel Low Alloyed asportato tramite tornitura
- Emissioni di CO₂ eq per ogni kg di Steel Low Alloyed asportato tramite fresatura
- Emissioni di CO₂ eq per ogni kg di Steel Low Alloyed asportato tramite foratura

Queste informazioni potranno in seguito essere ridotte tramite l'analisi di Pearson.

Queste informazioni potranno essere utilizzate per ottenere altre informazioni, oppure direttamente utilizzate per formare il database iniziale, che sarà utilizzato come input allo strumento di stima parametrica.

[Digitare qui]

4.4 Calcolo Emissione Di CO₂

Dopo aver raccolto le informazioni geometriche e quelle riguardanti le emissioni, si può procedere con il calcolo delle restanti informazioni per il completamento del input che sarà necessario in seguito.

Per ogni componente, ottenuto il volume asportato nella lavorazione, è stato possibile ottenere l'impatto ambientale [kg di CO₂ eq] tramite:

$$climate\ change = \left(\rho_k \cdot \sum_{i=1}^n V_{asp} \right) \cdot \delta_i$$

i, va da 1 a 3

Dove:

ρ_k , è la densità del materiale asportato

$\sum_{i=1}^n V_{asp}$, è la sommatoria dei volumi asportati in ciascun operazione

δ_i , emissione di CO₂ di ogni operazione

climate change = emissione di kg CO₂ eq. del componente

[Digitare qui]

4.5 Generazione Del Database Esteso

Dopo aver raccolto tutti i dati, è necessario procedere alla creazione del database esteso, il quale avviene tramite la moltiplicazione dei records, a cui corrispondono al numero dei componenti presi in esame.

La moltiplicazione dei record avviene sia tramite la scalatura dei modelli, sia l'inserimento dei diversi materiali.

Quindi ricordando che si hanno 73 modelli, e ogni modello verrà studiato per 3 diverse tipologie di materiale, si avranno 219 modelli. Invece per quanto riguarda i fattori di scalatura, ne sono stati assegnati due per ogni componente, quindi considerando che ogni componente ha sia la dimensione originaria che le due scalate, si avrà un totale di

$$219 \cdot 3 = 657$$

Dove 657 corrisponde al numero di modelli studiati,

4.5.1 Scalatura Dei Modelli

Questo procedimento consiste nell'assegnazione di fattori di scalatura alle dimensioni dei modelli 3D. Questi fattori sono dei valori positivi, che vengono moltiplicati per ogni parametro che definisce ognuno dei modelli presi in esame.

Una prima classificazione di tali fattori di scalatura, consiste nel definire se si parla di :

- Scalatura di ingrandimento, se il fattore è > 1
- Scalatura di riduzione, se il fattore è < 1

Un'ulteriore classificazione, può essere:

- Omogenea, se tutte le caratteristiche geometriche di un modello vengono moltiplicate per lo stesso fattore
- Eterogenea, se non tutte le caratteristiche geometriche di un modello vengono moltiplicate per lo stesso fattore.

[Digitare qui]

In questo studio è stata applicata una scalatura omogenea, quindi non è stata variata la morfologia rispetto ai modelli originari.

Il criterio applicato è stato quello di andare ad ingrandire modelli che erano particolarmente piccoli e andare ad rimpicciolire modelli che possedevano grandi dimensioni. Con tale metodologia è stato possibile avere un certo numero di modelli che ricoprì intervalli dimensionali piuttosto omogenei.

Per fare ciò, si sono suddivisi tutti i modelli in quattro sezioni in base alla loro massima direzione di estensione, e per ogni sezione si sono applicati due valori diversi del fattore di scala, ovvero:

- 0,05 -0,25 per la prima sezione di oggetti con maggior dimensione di estensione, modelli “grandi”
- 0,25-2 per la seconda sezione di oggetti “medio-grandi”
- 2-4 per la terza sezione di oggetti “medio-piccoli”
- 2,25-6 per la quarta sezione di oggetti “piccoli”

4.5.2 Inserimento Del Materiale

Per completare il database, è necessario tener conto della diversa tipologia di materiali studiati, che sono: Aluminium cast alloy, Chromium steel 18/8 e steel low alloyed.

Quindi da 219 modelli scalati , si arriva a 657 modelli totali.

Adesso si hanno a disposizione tutte le informazioni sufficienti per completare l’analisi

4.6 Analisi In RapidMiner

Una volta ottenuto il database finale, scremato tramite l’analisi di Pearson, comprensivo quindi del numero di modelli e del numero di parametri, si è potuto utilizzare lo strumento di stima RapidMiner.

Per l’analisi svolta, circa 80% dei modelli è stato utilizzato per la fase di training mentre circa il restante 20% è stato utilizzato per la fase di testing (percentuale che eventualmente può essere modificata).

4.6.1 Panoramica Generale

Dopo aver inserito il database completo e definito quale deve essere l'output, lo strumento mostra una prima panoramica riguardante tutte le diverse tipologie di analisi, che sono:

- *Deep Learning*: è una tecnica per apprendere esponendo le reti neurali ad ampie quantità di dati. Esso sfrutta diversi livelli di unità a cascata per svolgere compiti di estrazione di caratteristiche e di trasformazione. Risulta che ogni livello successivo utilizza l'uscita del livello precedente come input. Verrà trattata in seguito
- *Generalized Linear Model*: è un generalizzazione flessibile della regressione lineare ordinaria; dove quest'ultima riesce a prevedere il valore atteso di una quantità sconosciuta, come combinazione lineare di un insieme di valori osservati. Il Generalized Linear Model, generalizza la regressione lineare consentendo al modello lineare di essere correlato alla variabile di risposta tramite una funzione di collegamento.

- *Decision Tree*: è un modello predittivo, che può essere usato sia per casi di classificazione, sia per casi di regressione, questo vuol dire che l'output che andrà a prevedere sarà una variabile categorica o una quantità continua. Questo metodo di analisi, basa il suo funzionamento attraverso il continuo "spittamento" del input; un concetto fondamentale per capire il funzionamento del Decision Tree sono i "nodi", ovvero i luoghi dove i dati vengono splittati e le "foglie", ovvero i risultati intermedi o finali.

La logica che utilizza è il continuo cercare la condizione (nodi) che permette di generare le foglie in cui si distribuisce meglio le osservazioni del input

I vantaggi di tale metodologia è la velocità e la semplicità sia dal punto di vista dell'interpretazione che dal suo funzionamento, ma ha lo svantaggio che una piccola modifica del training può comportare una modifica significativa del risultato.

- *Random Forest*: anche questo può essere utilizzato sia per classificazione che regressione. Il Random Forest è un algoritmo "ensemble" ovvero un modello con al suo interno un insieme di altri modelli più semplici. Con il Random Forest, si andranno a creare tanti classificatori "deboli" che messi insieme forniscono un risultato ottimo.

La caratteristica di questo algoritmo è l'utilizzo di alberi indipendenti tra loro e di ugual importanza.

[Digitare qui]

I vantaggi sono la sua efficienza e la facilità di interpretazione ma ha lo svantaggio di avere tempo di esecuzione elevato.

- *Gradient Boosted Trees*:, è uno dei più utilizzati in quanto è di facile interpretazione e produce modelli accurati, verrà trattato nello specifico in seguito
- *Support Vector Machine*: è un modello di classificazione il cui obiettivo è quello di trovare la retta di separazione delle classi di training che massimizza il margine tra le classi stesse, dove con margine si intende la distanza minima della retta ai punti delle due classi di testing.

I vettori di supporto sono i valori di una classe più vicini alla retta di separazione, quindi quelli che si avvicinano di più all'altra classe e quindi quelli che si classificano con più maggiore difficoltà. (Figura 4.1)

Quello che differisce il Support Vector Machine è quello di classificare prendendo come riferimento gli esempi più difficili (i vettori di supporto) e non quelli più rappresentativi di una classe.

Maggiore sarà il margine, migliore sarà la generalizzazione; maggiore sarà il margine, maggiore è la distanza tra le classi e quindi minore la possibilità di fare confusione

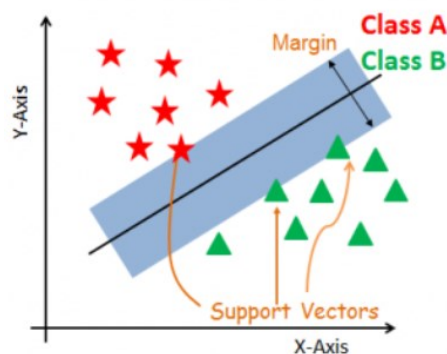


Figura 4 . 1 Support Vector Machine

[Digitare qui]

RapidMiner mostra un primo risultato confrontando tra di loro le diverse analisi, dando i seguenti risultati in termini di errore relativo e tempo di esecuzione dell'analisi, come si può vedere in *Figura 4.2* e *Figura 4.3*

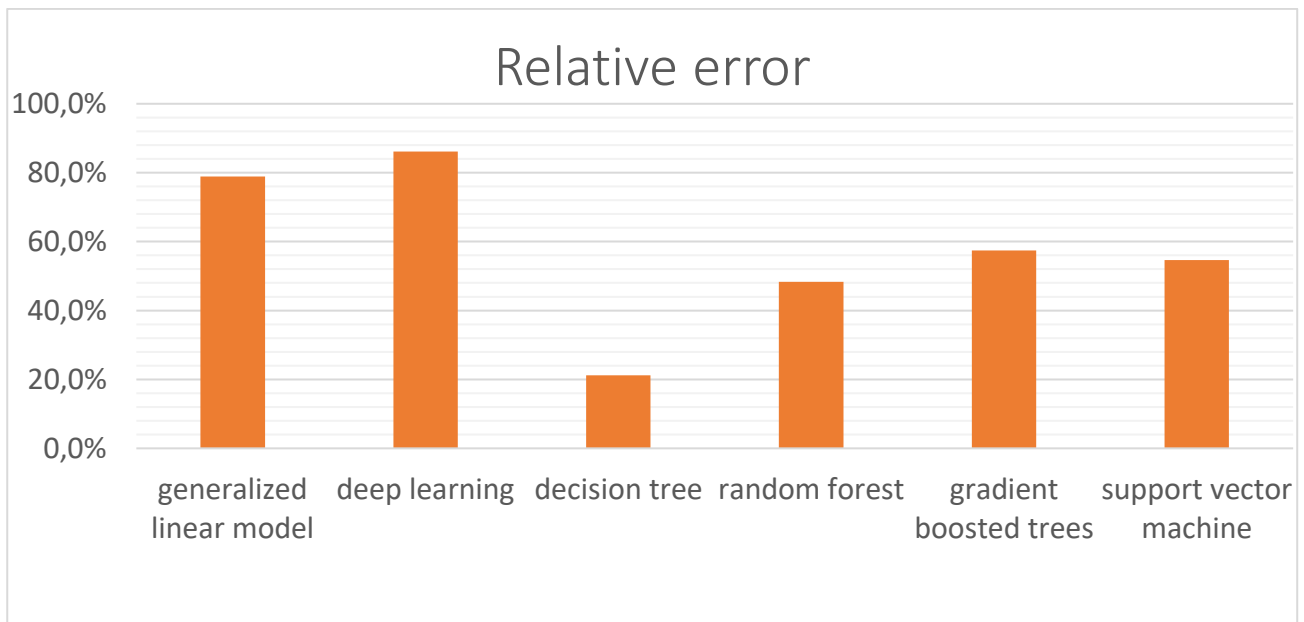


Figura 4 . 2 Errore relativo delle diverse analisi

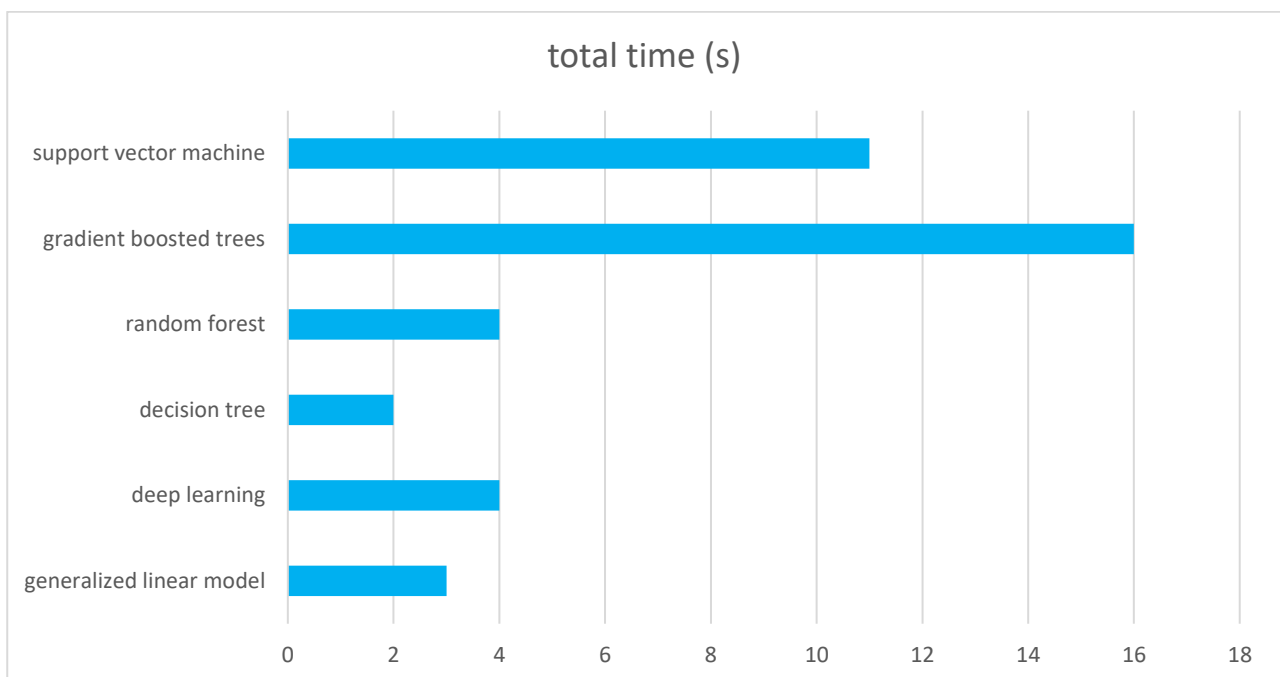


Figura 4 . 3 Tempo di esecuzione delle diverse analisi

[Digitare qui]

Come si vede da questa prima panoramica generale, si può notare come *il Deep Learning* sia quello che possiede un errore relativo maggiore invece *il Decision Tree* ha un errore relativo minore.

Per quanto riguarda invece in termini di tempo di esecuzione, si può notare come *il Gradient Boosted Trees* sia quello meno efficiente sotto questo punto di vista, invece anche sotto questo aspetto *il Decision Tree* sia il più efficiente.

Tra le tutte le tipologie di analisi sopra indicate, in questa tesi, si andranno a studiare *il Deep Learning* e *il Gradient Boosted Tree*.

4.7 Gradient Boosted Trees

Il Gradient Boosted Trees è definito come *ensemble*, ovvero un modello che contiene al suo interno modelli più semplici e questo comporta l'utilizzo di una combinazione di algoritmi più semplici al fine di ottenere prestazioni non raggiungibili utilizzandoli separatamente.

Il Gradient Boosted Trees (GBT), rappresenta una tecnica di regressione che costruisce un nuovo modello a partire da un'insieme di modelli predittivi deboli (*weak learner*) rappresentati da alberi di decisione.

Un albero decisionale (*Decision Tree*) è essenzialmente un layout di vari risultati associati a una serie di scelte legate l'una all'altra. *Il Decision Tree* è un modello predittivo che può essere usato sia per la classificazione, sia per la regressione.

L'albero decisionale, basa il suo funzionamento su il continuo "*splittamento*" dei dati andando a cercare di trovare il migliore condizione che permette di generare i risultati in cui si distribuiscono meglio le osservazioni del insieme di dati. Per capire il suo funzionamento bisogna definire per prima cosa, cosa sono i nodi e le foglie; i primi sono quelli dove avviene l'operazione di *split* dell'albero in più rami, invece le foglie sono i nodi che non vengono ulteriormente *splittati*, ovvero i risultati finali.

La costruzione di questi alberi decisionali implica quindi l'utilizzo di particolari metodologie per decidere sulla base di quali *feature* (caratteristiche) e sotto quali condizioni eseguire il processo di *split*, ma richiede anche di definire opportune condizioni di arresto.

In generale, è comunque presente alla base della costruzione, un algoritmo che si occupa di eseguire in maniera ricorrente una partizione delle features; ogni partizione è scelta selezionando, come già detto, il miglior *split* da un insieme di possibili *split*.

[Digitare qui]

Con il termine *Boosting*, si intende un algoritmo iterativo che definisce la modalità con cui combinare tra loro un insieme di modelli deboli (*weak learner*). Questo algoritmo segue una graduale costruzione, dove ogni step viene aggiunto un *weak learner* con lo scopo di migliorare i risultati ottenuti a partire dai modelli precedenti.

Il risultato finale, *strong learner*, sarà definito come la somma dei risultati dei *weak learner* costituenti, dove i *weak learner* in questo caso sono i *Decision Tree*, con lo scopo di andare a minimizzare la funzione che definisce il residuo tra la previsione reale e la previsione del modello.

Quindi lo scopo dell'algoritmo sarà quello di minimizzare il valore del residuo, cercando di tenderlo a zero, andando ad inserire continuamente nuovi *weal learner*.

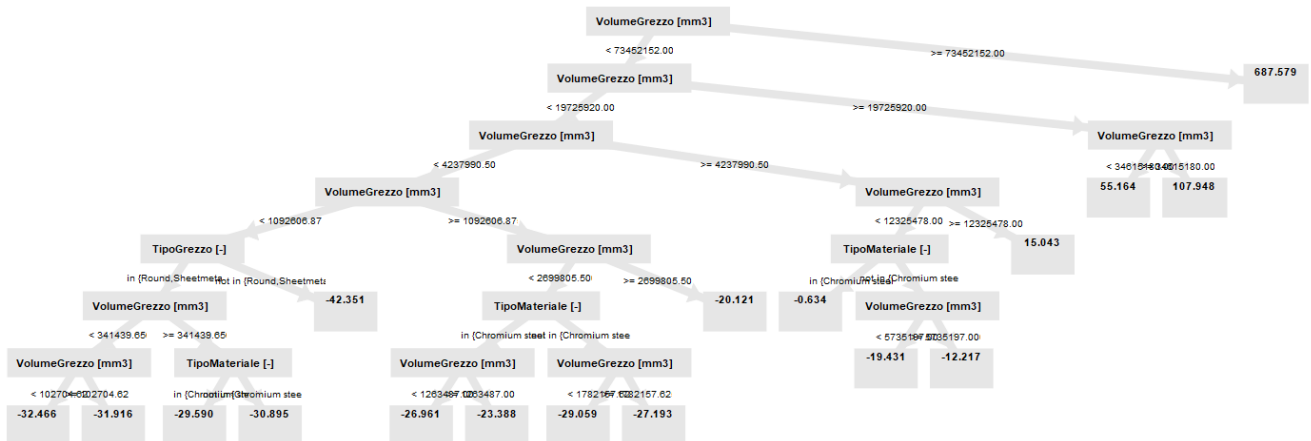


Figura 4 . 4 Decision Trees

La *Figura 4.4* sopra riportata, è stata ottenuta da RapidMiner per il caso trattato in esame, ed è importante andare a notare, quali sono i *Decision Trees* scelti dal algoritmo.

Dall'analisi *Gradient Boosters Trees*, tramite la funzione *Weights* su RapidMiner è possibile andare a valutare quali sono i parametri più influenti rispetto al output, ed da come si può vedere dalla *Figura 4.5*:

[Digitare qui]

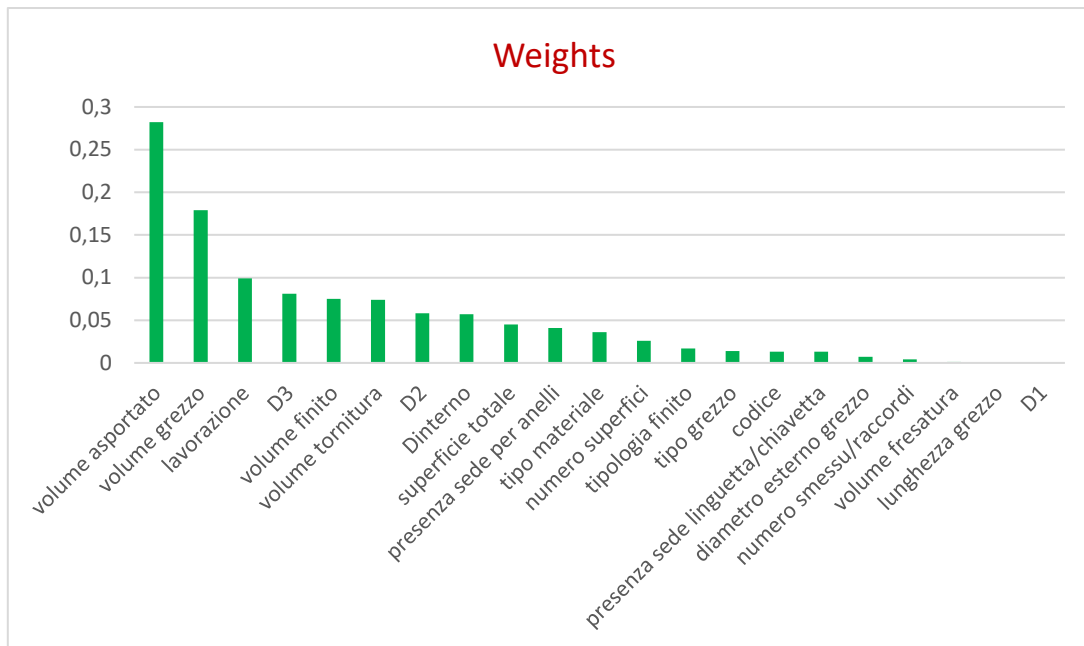


Figura 4 . 5 Gradient Boosted Trees Weights

Come è possibile notare dalla *Figura 4.5* i parametri più influenti secondo il *Gradient Boosted Trees*, sono:

- Volume asportato
- Volume grezzo
- Lavorazione
- Estensione assiale

Allo stesso modo si può notare come parametri la lunghezza del grezzo e l'estensione radiale non risultano influenti al fine della predizione dell'impatto ambientale e si sarebbe potuto arrivare a tale conclusione anche tramite l'analisi di Pearson.

Tramite il *Predictions Chart* è possibile osservare quanto si discosta la predizione dal valore effettivo, ovvero come è riportato in *figura 4.6*.

Dove si può vedere come la predizione sia tanto più accurata quanto è minore il valore di kg di CO₂ eq che viene emesso.

[Digitare qui]

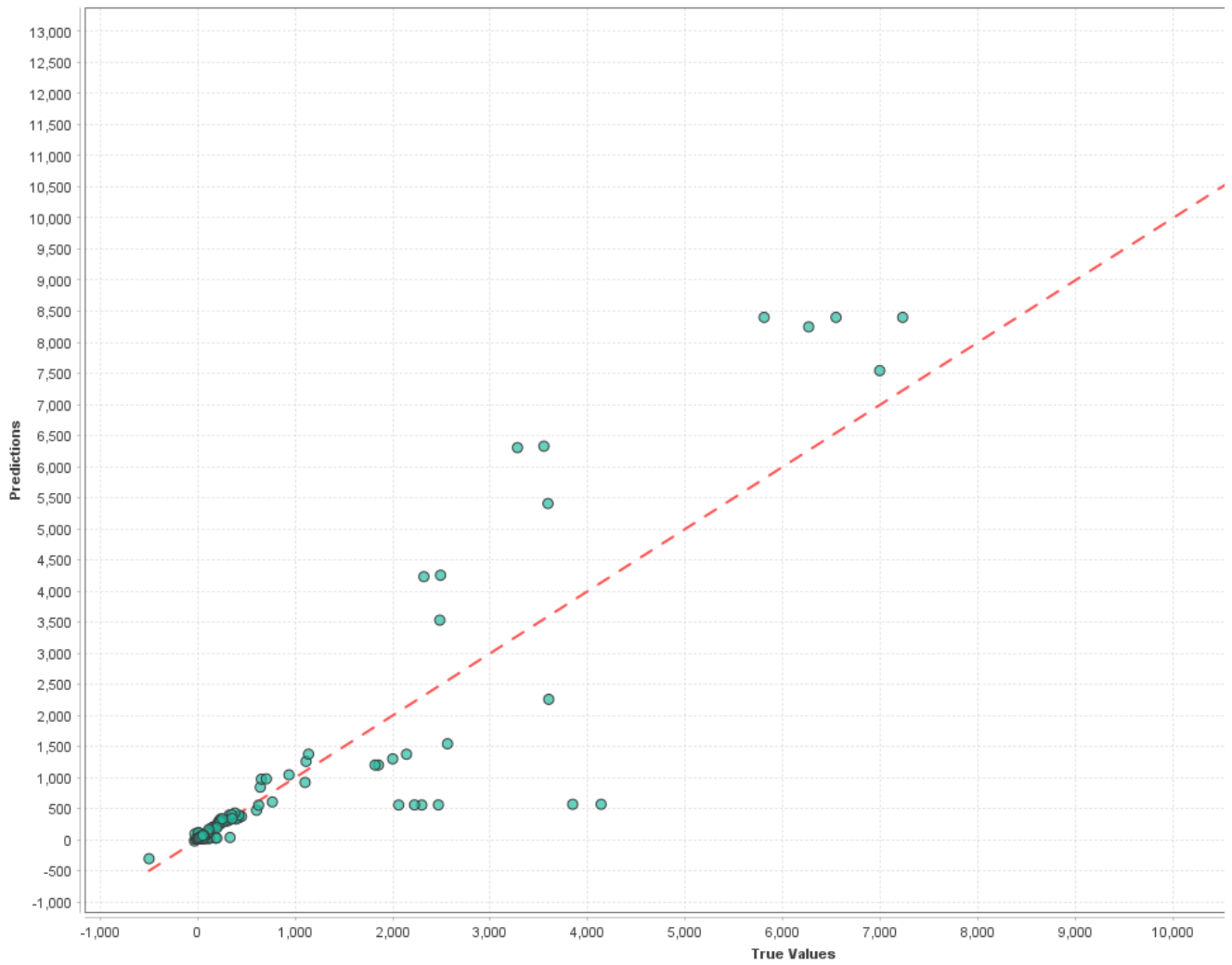


Figura 4 . 6 Predictions Chart Gradient Boosted Trees

[Digitare qui]

4.8 Deep Learning

Il *Deep Learning* realizza l'apprendimento avvelandosi di una rete neurale artificiale. Sfruttando la peculiare organizzazione dell'elaborazione nota come rete neurale, il *Deep learning* realizza l'apprendimento avvalendosi di classi specifiche di algoritmi disposti in più livelli di unità.

Questi livelli di unità, risultano essere non lineari e disposti a cascata, tale organizzazione permette al sistema di effettuare da sé le operazioni di: estrazione di caratteristiche, correlazione e organizzazione gerarchica dei dati con cui è alimentato. Ciascun livello successivo utilizza l'uscita del livello precedente come input.

Quindi il *Deep Learning* basa il suo funzionamento tramite le reti neurali artificiali multilivello, dove ogni livello è costituito da un ingresso (input), un livello nascosto (hidden) in cui avviene l'elaborazione dei dati, e un livello di uscita (output).

Risulta quindi che maggiore è il numero di livelli intermedi (hidden), maggiore sarà l'efficienza del risultato.

Ogni neurone artificiale presente in ogni strato costituisce un nodo computazionale dotato di connessioni verso altri neuroni e di parametri (pesi) attribuiti alle singole connessioni .

Per sviluppare queste reti neurali, si utilizzano linguaggi di programmazione come:

- Python
- Java

A differenza con i software tradizionali, però le reti neurali sono concepite per essere autodidatte nel comportarsi con i dati.

Solitamente quindi l'unica accortezza è nel "etichettare" i dati, ad esempio tramite il *meta tag*, ovvero parole chiave che permettono di eseguire l'analisi.

Questi sistemi, si basano su un processo di apprendimento "*trial and error*" e quindi affinché l'output sia affidabile, sono necessarie grandi quantità di dati; non solo, tali sistemi risultano difficili da governare a causa del elevato numero di livelli della rete neurale. Il numero di livelli e collegamenti tra neuroni nella rete è tale da poter rendere difficoltoso il calcolo delle regolazioni che devono essere apportate in ogni fase del processo; questo perché per il training si usano gli algoritmi di propagazione dell'errore, attraverso i quali si rivedono i pesi della rete neurale. In caso di errori, la rete emette all'indietro l'errore in modo che i pesi delle connessioni vengono modificati e aggiornati in maniera più opportuna; tutto questo in maniera iterativa finché l'errore non è nullo.

[Digitare qui]

Tramite il *Deep Learning*, avremo quindi a disposizione un ‘meccanismo’ che riesce a classificare in maniera autonoma i dati e a strutturarli in maniera gerarchica, andando a trovare quelli più utili alla risoluzione del problema, andando sempre a migliorarsi continuamente.

Di seguito verrà mostrata l’analisi tramite *Deep Learning* per quanto riguarda l’argomento trattato in questo studio; dove tramite la funzione *Weights*, come nel caso del *Gradient Boosted Trees*, è possibile andare a valutare quali sono stati i parametri più influenti rispetto al nostro output, ovvero come mostrato in *Figura 4.7* :

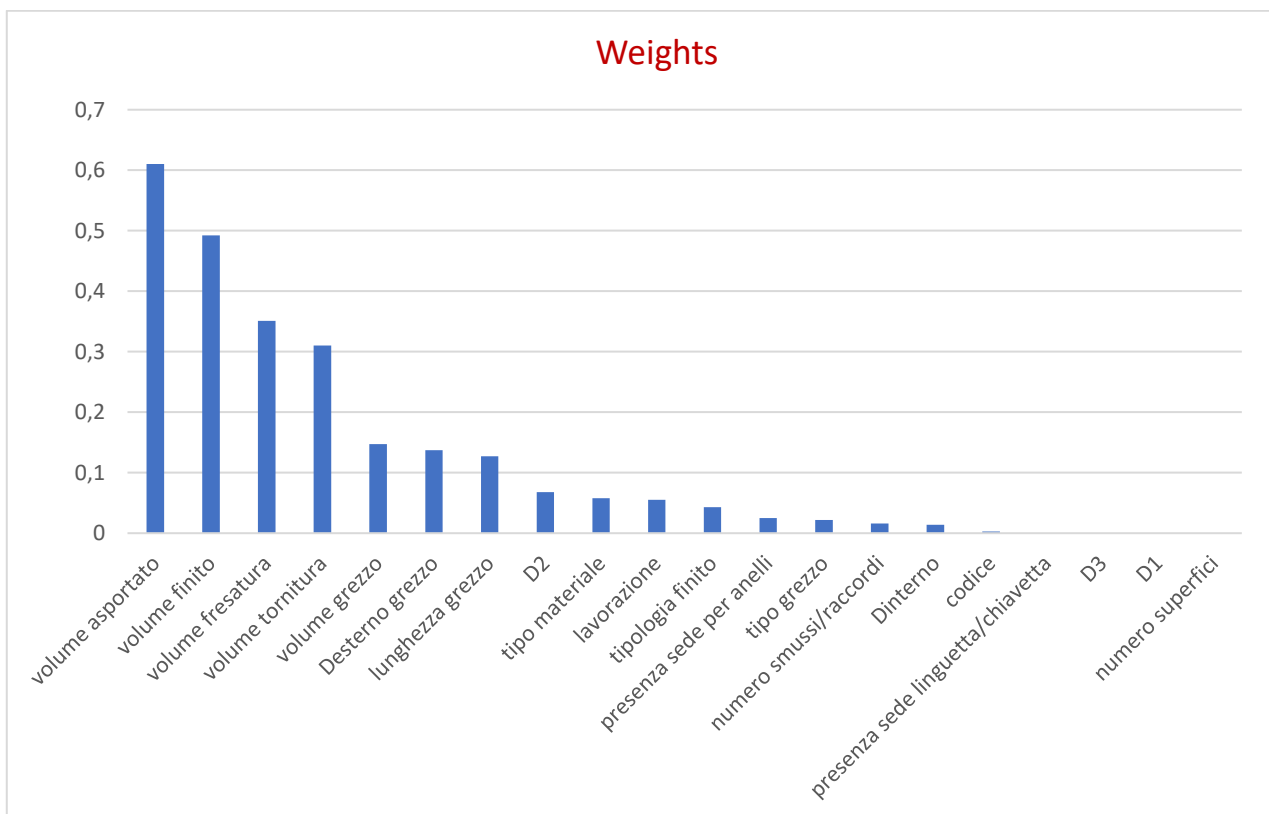


Figura 4.7 Deep Learning Weights

Come è possibile notare dalla *Figura 4.7*, i parametri secondo l’analisi tramite *Deep Learning*, che risultano maggiormente significativi sono:

- Volume asportato
- Volume finito
- Volume fresatura
- Volume tornitura

[Digitare qui]

Invece risultano essere presenti parametri che non influenzano in nessun modo l'emissione di anidride carbonica che sono :

- La presenza di eventuali sedi per linguetta/chiavetta
- Numero superfici
- Estensione in direzione assiale (D₃)
- Estensione in direzione radiale (D₁)

Il risultato che tali parametri non influenzano l'output, poteva essere ottenuto anche andando ad eseguire l'analisi di Pearson.

Come è stato fatto per l'analisi tramite *Gradient Boosted Trees*, anche per l'analisi tramite *Deep Learning* è possibile andare a valutare di quanto si discosta la predizione di emissioni di anidride carbonica dal valore effettivo, sempre tramite la funzione *Predictions Chart*, si ottiene la *Figura 4.8*

Da come si può vedere dalla *Figura 4.8*, l'analisi tramite *Deep Learning*, risulta essere maggiormente accurata per bassi valori di kg di CO₂ eq emessi.

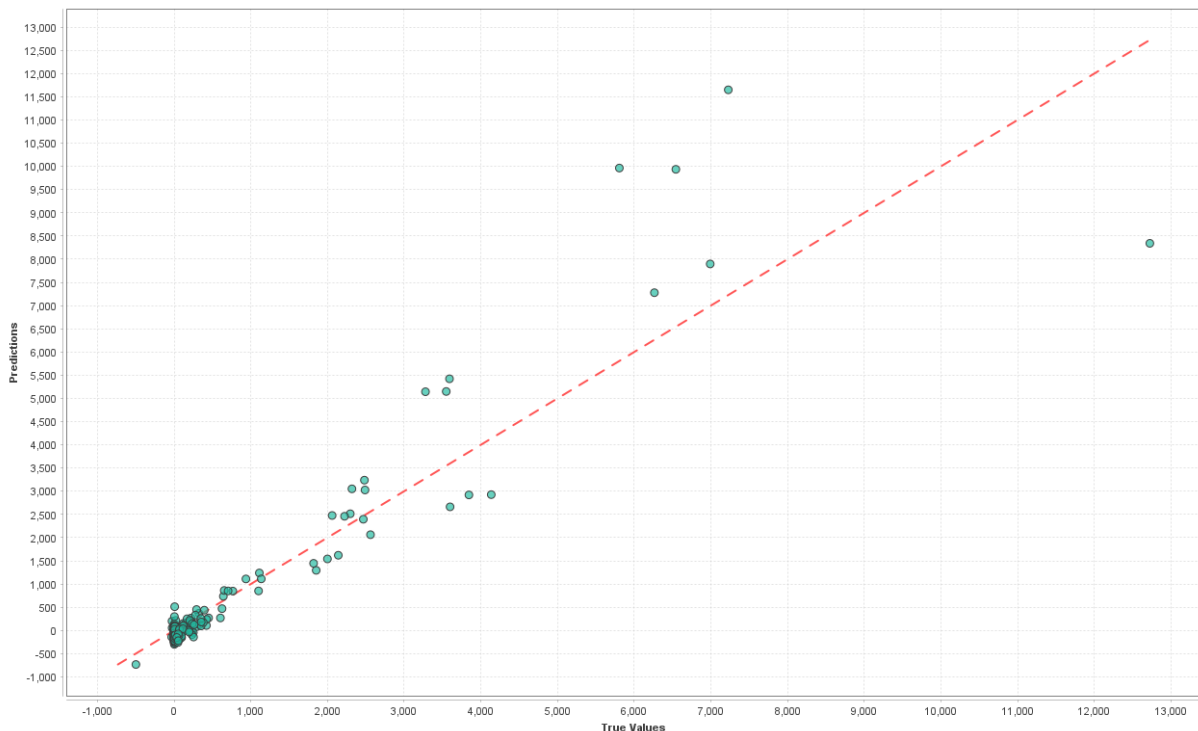


Figura 4 . 8 Deep Learning Predictions Chart

[Digitare qui]

4.8.1 Confronto Scelta Dei Parametri

In questo studio si sono andati a confrontare il *Deep Learning* e il *Gradient Boosted Trees*, dove il primo, realizza la previsione avvalendosi di una rete neurale artificiale multilivello, dove ogni livello è formato da un ingresso “*input*”, un livello nascosto “*hidden*” e un livello di uscita “*output*”; e tramite algoritmi di propagazione dell’errore riescono a filtrare i pesi delle connessioni tra i vari livelli e ottenere la migliore risposta. Per quanto riguarda invece il *Gradient Boosted Trees*, esso è formato da un insieme di algoritmi semplici (weak learner), rappresentati da alberi decisioni, e combinandoli tra loro fino ad arrivare soluzioni non raggiungibili considerandoli separatamente. Il suo funzionamento è basato sul continuo splittamento dei dati andando a cercare la migliore configurazione.

Tra il *Gradient Boosted Trees* e il *Deep Learning*, risulta che quest’ultimo sia meno preciso rispetto all’altro, in quanto va ad eseguire una previsione simile a una media.

Per quanto riguarda invece il *Gradient Boosted Trees*, risulta essere più preciso ma va ad assegnare i risultati per nuove casistiche non presenti nel database; risulta inoltre essere più economico rispetto all’altro.

In questo paragrafo si andrà a vedere nel dettaglio la differenza di risultato tra le due analisi, nello specifico, si andrà a valutare secondo ciascuna analisi quali sono le proprietà più influenti rispetto al nostro output.

Come si può vedere tramite *Figura 4.9*:

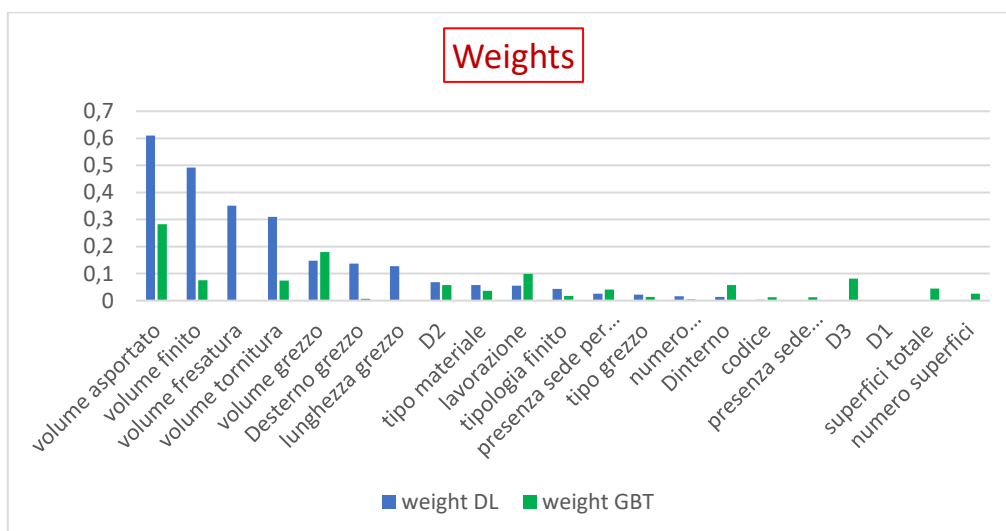


Figura 4.9 Deep Learning Weights

[Digitare qui]

Dalla Figura 4.9 , è possibile notare come sia molto diversa la scelta dei parametri per le due tipologie. Oltre alla diversa scelta dei parametri, balza all'occhio anche il valore del Weights effettivo tra le due analisi.

Per entrambe le analisi, è possibile vedere come il *volume asportato*, sia il parametro più influente rispetto alle emissioni di carbonio, in quanto maggiore sarà il volume da asportare, maggiore sarà l'energia necessaria e quindi maggior consumo di combustibili fossili e quindi questo risultato risulta essere in accordo con una casistica reale.

4.9 Esempi Applicativi

In questo capitolo, si andranno a valutare e confrontare i risultati ottenuti tramite l'analisi con RapidMiner per tre diversi componenti e ognuno dei quali nei tre diversi materiali trattati.

I componenti selezionati per andare a eseguire questo confronto sono:

- 1) T0001670 (*Figura 4.10*)

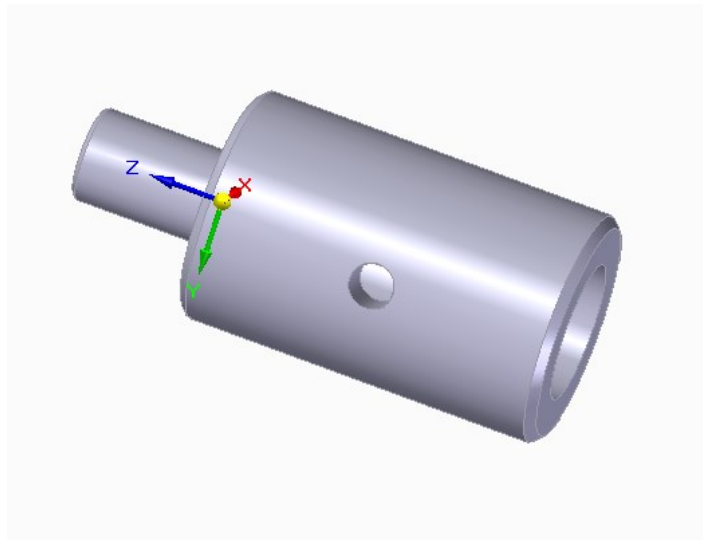


Figura 4 . 10 Componente T001670

[Digitare qui]

2) T0042708 (Figura 4.11)

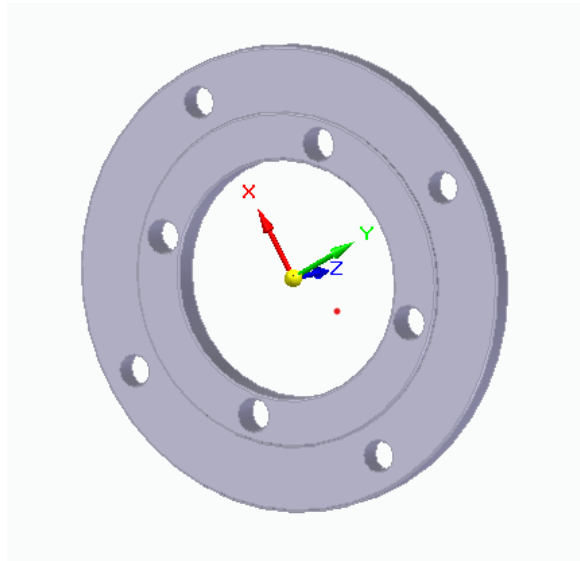


Figura 4 . 11 Componente T0042708

3) Albero (9) (Figura 4.12)

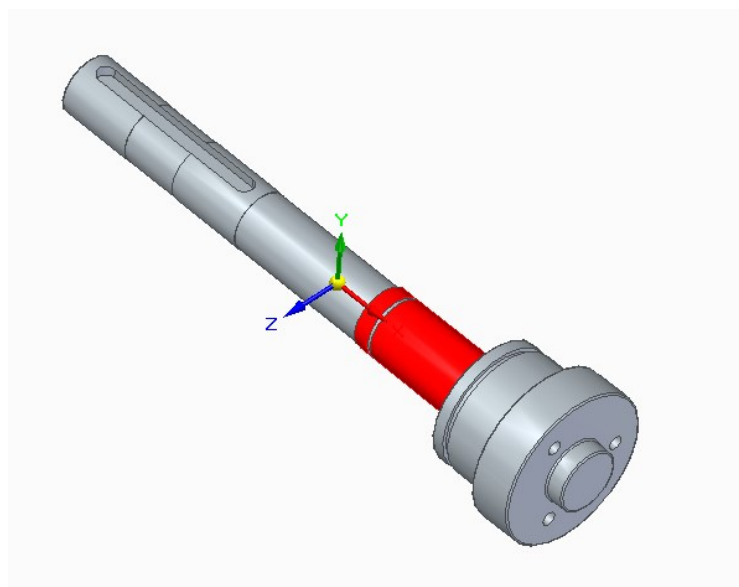


Figura 4 . 12 Componente "Albero (9)"

Per ciascuno dei componenti si è andati a valutare l'errore relativo (E_r) tra il valore predetto da quello effettivo, ovvero :

$$E_a = |y_i - \hat{y}_i|$$

[Digitare qui]

$$E_r = \left(\frac{E_a}{V_m} \right) \cdot 100$$

Dove:

y_i è il valore predetto

\hat{y}_i è il valore previsto

V_m è il valore medio delle previsioni, in questo caso coincide con y_i

E_a è l'errore assoluto

E_r è l'errore relativo, in percentuale

I risultati ottenuti sono rappresentati dalla *Tabella 4*, dove si può notare anche attraverso il valore di errore relativo il *Gradient Boosted Trees* sia più preciso rispetto al *Deep Learning*.

Tabella 4 Errore delle Predizioni di Impatto Ambientale

COMPONENTE	MATERIALE	CLIMATE CHANGE [kgCO2 eq]	PREDICTION DEEP LEARNING [kgCO2 eq]	PREDICTION GRADIENT BOOSTED TREES [kgCO2 eq]	E_{rDP} [%]	E_{rGBT} [%]
T00016170	STEEL LOW ALLOYED	8,4	133,3	28,4	93,6	70,2
	CHRONIUM STEEL 18/8	295	82,1	299,7	259,4	1,5
	ALLUMINIUM CAST ALLOY	9,2	81,2	30	88,6	69,3
T00042708	STEEL LOW ALLOYED	4,3	67,6	34,5	93,6	87,4
	CHRONIUM STEEL 18/8	8,2	99,7	114,1	91,7	92,8
	ALLUMINIUM CAST ALLOY	35,1	151,1	48,1	76,7	27
ALBERO (9)	STEEL LOW ALLOYED	151,2	198,2	155,6	23,7	2,8
	CHRONIUM STEEL 18/8	278	330,4	310,1	15,8	10,3
	ALLUMINIUM CAST ALLOY	6263,5	8248,1	7278,1	24	13,9

[Digitare qui]

Andando a mediare i valori di errore relativo riguardante il *Deep Learning* (E_{rDP}), si ottiene il valore di 86%, in accordo con il valore previsto all'inizio dell'analisi, come mostrato in *Figura 4.2* .

L'elevato errore è dovuto al fatto che il *Deep Learning* si basa su un processo di approfondimento *'trial and error'* e quindi affinché l'output sia affidabile ha bisogno di un enorme quantità di dati.

[Digitare qui]

5 CONCLUSIONE

L'obiettivo di questo studio era quello di determinare un modello parametrico basato su tecniche di machine learning che fosse capace a determinare l'impatto ambientale di componenti assialsimmetrici.

Per determinare un modello che sia capace di determinare quanto inizialmente prefissato, è necessario fornirgli dei dati di input.

Al tal fine, per determinare questi dati, si è utilizzato come primo strumento, il software SimaPro, il quale ha permesso di ricavare dati riguardanti le emissioni di CO₂ eq. per ciascuno dei materiali studiati e per le rispettive lavorazioni che ciascun componente andrà a subire.

Questi dati riguardanti le emissioni, andranno a essere combinati insieme ai dati geometrici dei componenti ricavati tramite l'analisi CAD ed opportunamente scalati in maniera omogenea andando così a ricoprire un campo di studio più ampio.

Una volta ottenuto l'input finale, è possibile andare a determinare tramite RapidMiner , il modello parametrico tramite due diversi algoritmi, ovvero:

- *Deep Learning*
- *Gradient Boosted Trees*

Da queste due diverse analisi è stato possibile notare come quella tramite *Deep Learning* risulta avere un errore maggiore nella predizione di emissioni di CO₂.

Il miglioramento dei risultati ottenuti può avvenire tramite un maggiore numero di componenti, un maggiore numero di componenti aventi geometrie più diverse possibili e complesse, aggiungendo parametri che possono risultare rappresentativi ed efficaci (come per esempio il "mix elettrico nazionale" del sito produttivo trattato in questione, che corrisponde all'insieme delle soluzioni tecnologiche e l'insieme delle fonti energetiche impiegate per produrre energia elettrica che quotidianamente utilizza una nazione), ed infine effettuare una più ampia moltiplicazione del modello.

[Digitare qui]

6. BIBLIOGRAFIA

Bin Song, B. S. (2018). *Data-driven Approach for Discovery of Energy Saving Potentials in Manufacturing Factory*. Singapore.

Erik Hofmann, M. R. (2017). *Industry 4.0 and the current status as well as future prospects on logistics, Volume 89*.

Gro, B. (1987). *rapporto Brundtland*.

ISO. (2006). *1404:2006 - Environmental management - LCA - Principles and Framework*.

ISO. (2006). *Environmental management- LCA- Requirements and Guidelines*.

Katili AR, B. R. (2015). *Space Cooling in Buildings in Hot and Humid Climates - a Review of the Effect of Humidity on the Applicability of Existing Cooling Techniques*. Nottingham.

krivosheeva OI, M. M. (2022). *The economic effect of public–private partnerships in the implementation of climate projects for the disposal of municipal solid waste*.

McGinty, L., & Smyth, B. (2000). *In Proceedings of the European Workshop on Advances in Case-Based Reasoning*. Trento .

Patrick Schneider, F. X. (2022). *Anomaly Detection and Complex Event Processing over IoT Data Streams* .

Pfeiffer, S. (2017). *La visione di "Industria 4.0" in divenire: un caso di futuro raccontato, addomesticato e scambiato*. Nanoetica 11 .

Pozzolo, P. (2020). *La correlazione lineare di Pearson*.

Tu Peng, X. Y. (2020). *Constructing an Environmental Friendly Low-Carbon-Emission Intelligent Transportation System Based on Big Data and Machine Learning Methods*. Beijing.