



UNIVERSITÀ POLITECNICA DELLE
MARCHE

FACOLTÀ DI INGEGNERIA

Corso di laurea triennale in Ingegneria Meccanica

**Misure innovative per il controllo di ambienti
indoor**

Innovative measures for the control of indoor environments

Relatore: **Prof. Ing. Gian Marco Revel**

Laureando: **Matteo Vallasciani**

Correlatore: **Sara Casaccia**

Correlatore: **Nicole Morresi**

INDICE

1. INTRODUZIONE	4
2. STATO DELL'ARTE	5
2.1. Il comfort termico	5
2.2. Metodologie innovative di controllo del comfort termico	6
3. MATERIALI E METODI	16
3.1. Test preliminare	16
3.2. Test con termocamera	19
3.2.1. Descrizione del progetto.....	19
3.2.2. Set-up sperimentale	19
3.2.3. Sensori utilizzati e acquisizione dati	22
3.3. Analisi delle immagini termografiche.....	24
3.4. Analisi dei dati	29
3.4.1. Machine Learning Supervisionato.....	29
3.4.2. Logistic Regression	31
3.4.3. Algoritmo per l'analisi dei dati	34
3.4.4. Performance dell'algoritmo.....	36
4. RISULTATI	37
4.1. Risultati del test preliminare	37
4.2. Risultati del test con la termocamera	42
5. CONCLUSIONI	47
6. BIBLIOGRAFIA	49

1. INTRODUZIONE

Il lavoro presentato in questa tesi di laurea si inserisce all'interno di un più ampio progetto europeo portato avanti dai ricercatori del Dipartimento di Ingegneria Industriale e Scienze Matematiche (DIISM) dell'Università Politecnica delle Marche in collaborazione con altre università europee.

Per prima cosa è stata approfondita la parte letteraria nella quale si sono valutati diversi sistemi innovativi sviluppati da ricercatori di tutto il mondo per il controllo del comfort termico umano in ambienti indoor. Il problema del benessere termico degli esseri umani è infatti molto rilevante e difficilmente controllabile a causa della sua natura strettamente soggettiva. L'altra problematica fondamentale legata ad esso è l'eccessivo consumo di energia richiesto dai sistemi HVAC centralizzati per il suo mantenimento.

Il lavoro di tesi presentato si è focalizzato sullo studio di diversi parametri che sono correlati alla sensazione termica dell'utente, con l'obiettivo di individuare quello che permette in maniera migliore di prevederla e di adottare un metodo di controllo ottimale. In una prima fase del lavoro, è stato valutato l'impatto dei parametri ambientali quali la temperatura, l'umidità e velocità dell'aria, mentre in una seconda fase ci si è concentrati sulla misura dei parametri fisiologici quali la temperatura della fronte e delle guance, per cercare di stabilire una correlazione fra parametri fisiologici e sensazione termica dell'utente.

In seguito, è stato utilizzato un algoritmo di Machine Learning utilizzando il linguaggio di programmazione Python, tramite il quale è stato possibile ottenere delle previsioni della sensazione termica dei soggetti sottoposti alle prove effettuate, al variare di parametri ambientali e parametri fisiologici. Al termine del progetto si è poi verificata l'efficacia del modello predittivo sviluppato ed è stato possibile individuare quale dei vari parametri utilizzati permette una più accurata predizione del comfort termico dell'utente.

2. STATO DELL'ARTE

2.1. Il comfort termico

Il comfort termico di un essere umano è definito come la “condizione che esprime la soddisfazione con l’ambiente termico circostante. L’insoddisfazione può essere causata dal disagio (discomfort) del corpo umano nel suo insieme in relazione al caldo o al freddo, o dal raffreddamento (o riscaldamento) di una particolare parte del corpo” [UNI EN ISO 7730]. La condizione di comfort coincide dunque con lo stato in cui il soggetto non sente caldo né freddo ed è stimato a partire da una valutazione soggettiva [ASHRAE 55 – 2004]. Per questo motivo riuscire a definire il concetto di comfort risulta complicato perché esso non è meramente legato ad una evidenza scientifica ma è piuttosto una preferenza personale/sociale/circostanziale [1]. Sulla percezione del comfort incidono dunque numerose variabili legate sia all’ambiente circostante sia a fattori personali dell’essere umano. La norma UNI EN ISO 7730 individua 6 fattori principali che influenzano il comfort termico; 4 legati ai parametri ambientali e 2 legati a quelli personali (fig.1).

- Temperatura dell’aria
- Umidità relativa
- Velocità dell’aria
- Temperatura media radiante
- Valore dell’isolamento del vestiario
- Tasso metabolico



Figura 1- Comfort termico e i parametri che lo influenzano.

Per la valutazione del comfort termico esistono diversi modelli basati sia su aspetti fisiologici e comportamentali sia su analisi statistiche, in modo da coprire anche le diverse sensibilità dei soggetti. I modelli più utilizzati sono due: il modello di Fanger e il modello adattivo. Il primo è stato sviluppato negli anni '60 ed è lo standard a cui si rifanno le principali norme, tra cui la UNI EN ISO 7730. Esso

deriva da studi statistici condotti in laboratorio ed è basato sull'idea che il corpo umano sia un sistema termodinamico che scambia calore e lavoro con l'ambiente esterno [13]. L'altro modello, quello adattivo, è più recente e deriva da studi statistici condotti in edifici reali. È meno restrittivo sulla soddisfazione termica del soggetto rispetto al modello Fanger e permette di ottenere comfort a costi energetici inferiori [13].

L'indice per la valutazione del comfort termico attualmente più utilizzato è il Predicted Mean Vote (PMV), proposto da Fanger e adottato dalla norma UNI EN ISO 7730 [13]. Esso è funzione delle 6 variabili già incontrate ed è definito su di una scala di 7 valori: da -3 (freddo) a +3 (caldo) dove il valore 0 corrisponde alla condizione neutrale di comfort. Un altro indice molto utilizzato è la percentuale di persone insoddisfatte (PPD). La norma UNI EN ISO 7730 definisce l'intervallo di comfort termico per valori di PMV compresi tra -0.5 e +0.5 e un valore massimo del PPD pari al 10 % [13].

2.2. Metodologie innovative di controllo del comfort termico

Gran parte degli studi sul comfort termico si sono focalizzati nel cercare di massimizzare il numero di persone in condizione di benessere termico in un ambiente indoor.

Il caso più studiato è quello relativo all'ambiente lavorativo in cui si è notato come un ambiente termicamente soddisfacente può effettivamente migliorare la produttività dei lavoratori stessi. Secondo lo standard ASHRAE 55, i range di temperatura in un edificio di tipo lavorativo devono essere tra i 20 °C e i 24 °C nella stagione invernale e tra i 24 °C e i 27 °C in quella estiva [2]. All'interno di questi range di temperatura però, essendo la percezione del comfort termico fortemente soggettiva, alcuni soggetti potrebbero comunque trovarsi in condizione di discomfort.

In aggiunta si è notato come i consumi energetici per il mantenimento del comfort termico sono molto elevati. Ad esempio, nelle nazioni più sviluppate si aggirano tra il 20% e il 40% dei consumi annui e, nonostante ciò, vi è comunque un'elevata percentuale di lavoratori che riscontra una condizione di stress termico. Nasce così l'esigenza di superare l'utilizzo di sistemi di condizionamento centralizzati, favorendo l'uso di sistemi di comfort termico personalizzati per ogni individuo, con l'obiettivo dunque di aumentare la percentuale di soggetti nella condizione di benessere e, allo stesso tempo, di andare a ridurre gli elevati consumi di energia.

Si basano su questi obiettivi diversi studi improntati sulla ricerca di nuove metodologie di rilevamento e controllo del comfort termico. Per metodologia di controllo si intende una strategia capace di garantire il comfort termico degli utenti. Essa al suo interno comprende tutti quei componenti che vengono utilizzati per la raccolta dei dati del soggetto e la loro rielaborazione e, tramite di essi, riesce a controllare la percezione termica dell'individuo grazie a dispositivi fisici attui al controllo dei parametri dell'ambiente indoor (ad esempio tramite termostati).

L'articolo "Development and validation of a smart HVAC control system for multi-occupant offices by using occupants' physiological signals from wristband" mira a sviluppare e convalidare una strategia di controllo per i sistemi HVAC in uffici condivisi da più soggetti [3]. L'obiettivo è quello di andare a sviluppare un modello capace di integrare i parametri ambientali e i parametri fisiologici, così da ottenere un miglior controllo del comfort termico.

I test sono stati effettuati in 7 uffici condivisi nell'Università di Purdue, Stati Uniti, su 24 studenti diversi (16 uomini e 8 donne). La durata dei test è stata di 40 minuti e si sono effettuati test con temperatura crescente (riscaldamento) e test con temperatura decrescente (raffreddamento), fig. 2.

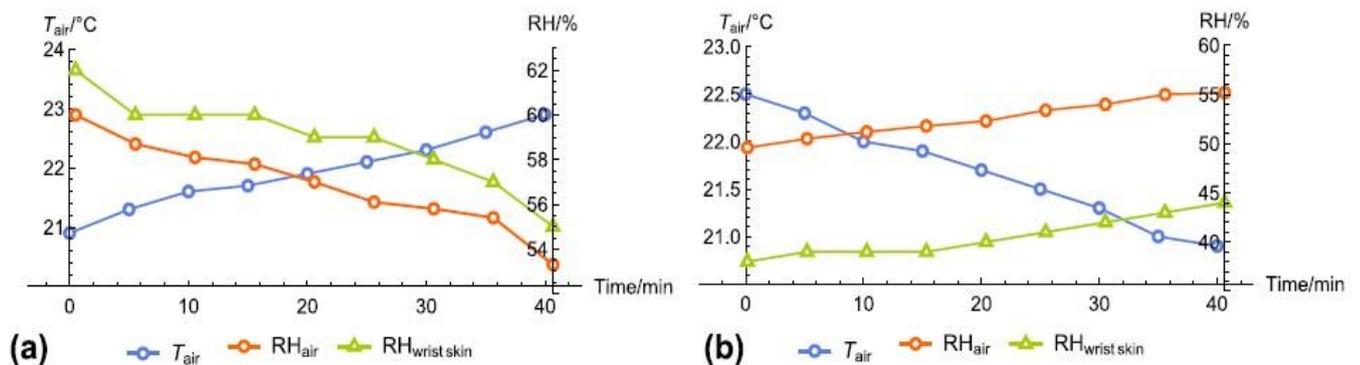


Figura 2 - Variazioni di temperatura, umidità dell'aria e umidità della pelle del polso in 40 minuti; (a) caso di riscaldamento, (b) caso di raffreddamento

Durante questi test sono stati raccolti dati ambientali, quali la temperatura e l'umidità dell'aria, e dati fisiologici, quali la temperatura e l'umidità della pelle del polso, l'Heart Rate (HR) e la resistenza termica del vestiario (CLO). La frequenza di campionamento dei dati è stata di 10 minuti e i parametri fisiologici sono stati rilevati tramite un *wristband*. Ogni 10 minuti è stato anche chiesto ai partecipanti di compilare il questionario relativo al Thermal Sensation Vote (TSV), nel quale era richiesto di esprimere un voto da -3 (eccessivamente freddo) a +3 (eccessivamente caldo) dove lo 0 rappresenta la condizione neutrale di comfort.

Una volta raccolti i dati, essi sono stati elaborati attraverso l'utilizzo di una rete neurale artificiale, Modello ANN, capace di elaborare delle relazioni tra le variabili di input (dati ambientali e fisiologici) e di output (TSV) e di prevedere poi il TSV dei soggetti. Per lo sviluppo del modello ANN si è utilizzato Matlab Deep Learning Toolbox.

Dopo aver allenato il modello ANN, si è sviluppata una strategia di controllo per il sistema HVAC utilizzando la correlazione tra i parametri fisiologici e il TSV degli utenti. In figura 3 si può osservare il principio di funzionamento della strategia di controllo.

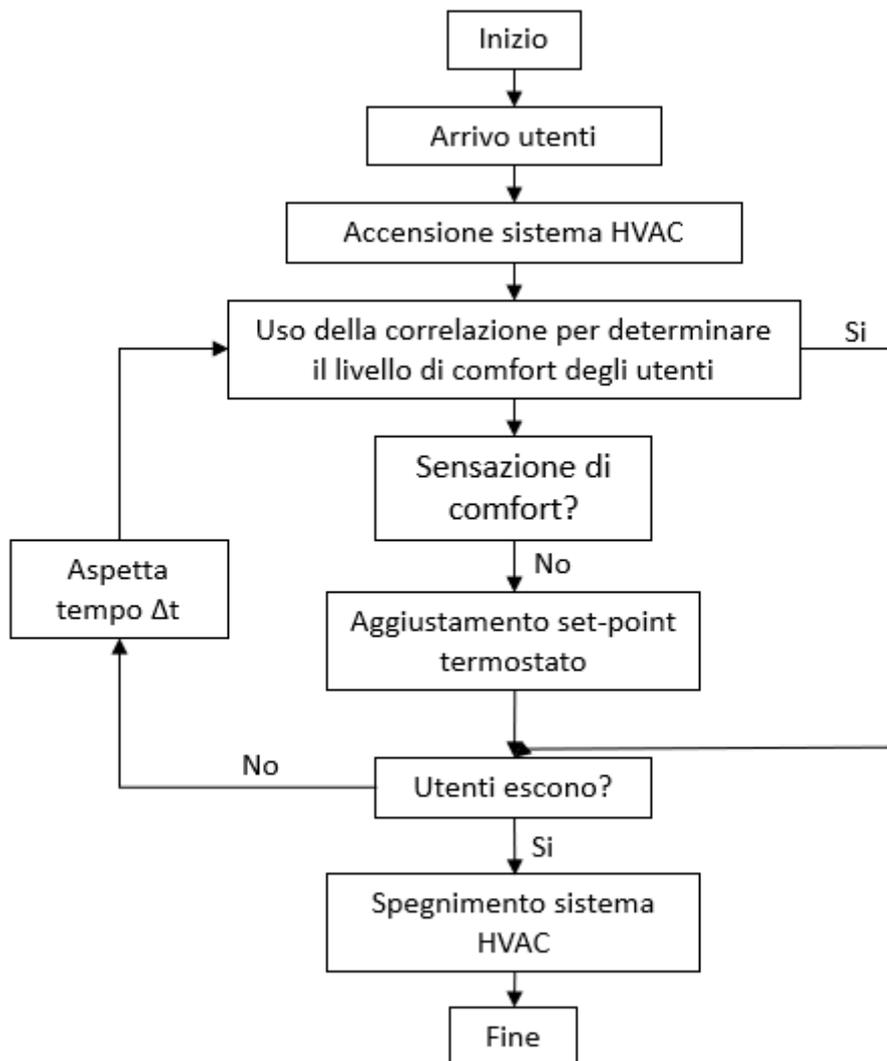


Figura 3 - Principio di funzionamento della strategia di controllo del sistema HVAC

Il sistema Bluetooth degli wristbands rileva la presenza degli utenti nella stanza e viene così attivato il sistema HVAC. Il braccialetto rileva i dati fisiologici ogni 10 minuti e tramite il modello ANN si ottiene una previsione del TSV e, sulla base di quest'ultima, si verifica la necessità o meno di

modificare il set-point del termostato. Quando gli utenti lasciano la stanza, il sistema HVAC viene spento, sempre grazie all'informazione ricevuta tramite il Bluetooth del braccialetto.

Il modello ANN è quindi in grado di predire il TSV degli utenti utilizzando 6 parametri di input. L'accuratezza del modello è maggiore nel caso in cui vengono integrati i parametri ambientali e fisiologici, con un R^2 pari a 0.89, mentre quando si escludono i parametri fisiologici si ha un R^2 uguale a 0.75.

Per quanto riguarda il sistema di controllo, esso permette di aumentare in maniera sensibile il numero di utenti in condizione di comfort (fig.4). Prima di utilizzare il sistema di controllo, oltre la metà dei soggetti si sentiva a disagio, variando da freddo (TSV = -2) a caldo (TSV = 2). Dopo aver utilizzato il sistema, quasi tutti gli occupanti riportano una sensazione neutra. Meno del 10% degli occupanti avverte ancora leggermente caldo o freddo, mentre nessuno degli utenti riscontra un eccessivo discomfort (TSV=|2|, TSV=|3|).

Dal punto di vista energetico, il consumo di energia del sistema HVAC utilizzando la strategia di controllo tramite il *wristband* è praticamente lo stesso di quello che si otterrebbe con un set-point costante.

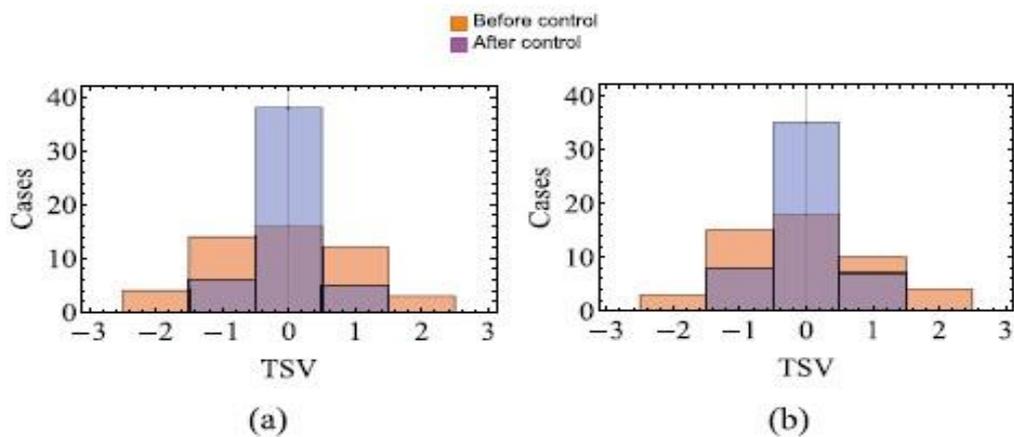


Figura 4 - Distribuzione del TSV degli utenti prima e dopo aver utilizzato il sistema di controllo.

Un altro articolo di interesse è “Personalized human comfort in indoor building environments under diverse conditioning modes” [4]. Anche in questo studio l'obiettivo è quello realizzare una previsione del TSV, e quindi del comfort degli utenti, integrando i dati ambientali e fisiologici. Si vuole quindi

sviluppare un framework capace di migliorare il comfort termico complessivo in ambienti indoor a condivisione singola e multipla.

I test relativi ad ambienti a occupazione singola sono stati effettuati nel Luglio 2016 in 3 camere distinte ad Ann Arbor (Michigan), ognuna dotata di un sistema HVAC e con una temperatura variabile tra 16.7 °C e 28.7 °C, sia in modalità di ventilazione naturale che forzata. I test relativi ad ambienti condivisi invece, si sono svolti a Novembre 2016 all'interno di un ufficio a Madison (Wisconsin). Si sono svolti per l'intera durata della giornata lavorativa, dalle 8 alle 18, ed erano presenti 20 impiegati. In questo caso si è stata utilizzata unicamente la modalità di condizionamento forzata.

Per entrambi i test sono stati rilevati parametri ambientali quali la temperatura, l'umidità e la CO₂ dell'aria, mentre per quanto riguarda i parametri fisiologici sono stati considerati la temperatura della pelle e l'HR, entrambi rilevati tramite una *Microsoft Band2*. I parametri ambientali quali la temperatura e l'umidità dell'aria invece sono stati rilevati tramite un sensore digitale *Sensorist Wireless Pro T/H* (precisione: ± 0.2 °C e $\pm 3\%$). Per la CO₂ si è invece utilizzata una sonda COZIR (precisione: ± 50 ppm). Il TSV anche in questo caso è stato ottenuto tramite la compilazione di un questionario.

Tutti i dati raccolti sono poi stati integrati all'interno di una Phone-App, collegata via bluetooth con la *Microsoft Band2*. L'applicazione scambia poi informazioni con un Database centrale, MongoDB, nel quale vengono archiviati i dati. A questo punto, tramite l'algoritmo di classificazione Random Forest, si ottengono delle previsioni del TSV, basate su correlazioni tra dati ambientali e fisiologici.

Si ha poi uno Script Python, composto da due diversi algoritmi: un algoritmo di *selezione* e un algoritmo di *decisione collettiva*. Si esegue prima quello di selezione, tramite il quale viene impostata una tra le modalità di condizionamento naturale e forzata, e successivamente quello di decisione. Quest'ultimo, attraverso una previsione del TSV, valuta il punteggio di comfort di gruppo per verificare se la scelta di condizionamento naturale/forzata selezionata con il primo algoritmo sia corretta o se occorre modificarla. Attraverso lo Script Python viene così programmato un NEST (termostato), capace di regolare il Set-point di temperatura massimizzando il livello di comfort complessivo. L'intero framework di controllo sviluppato è illustrato nella figura 5.

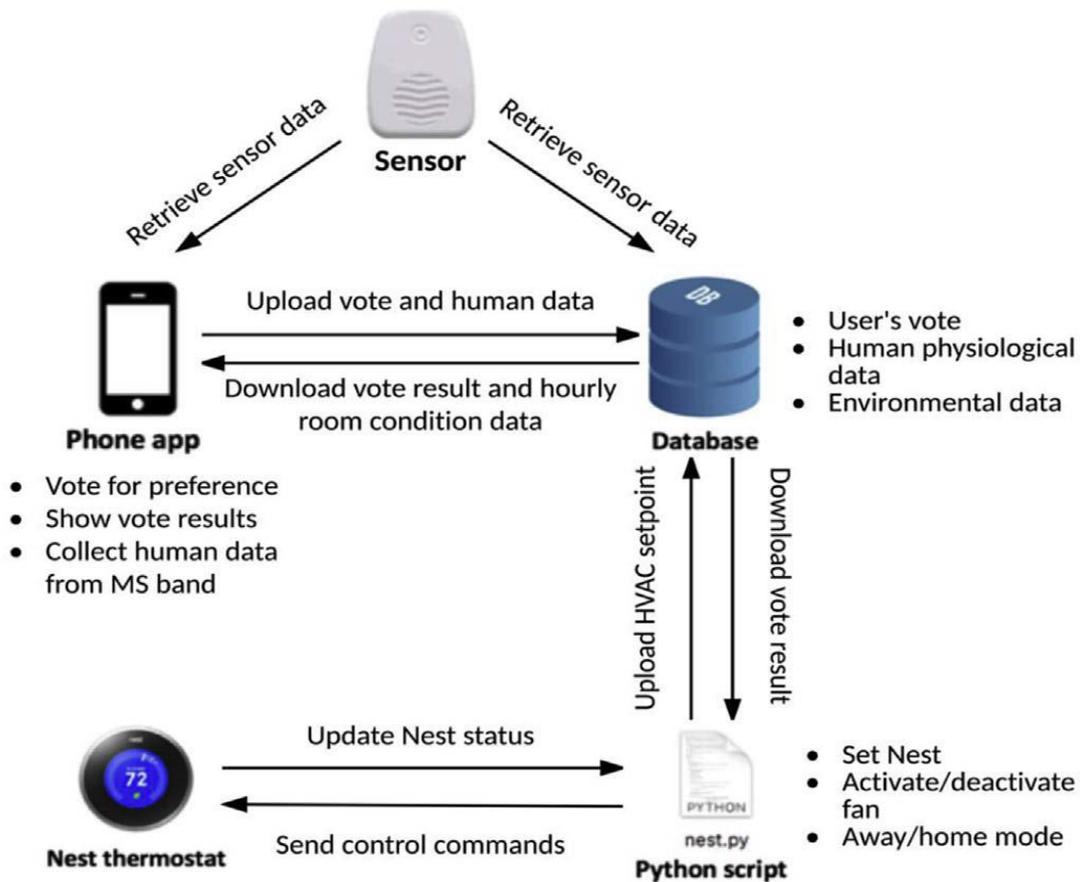


Figura 5 - Componenti del framework di controllo del sistema HVAC.

Il framework sviluppato in questo studio registra l'80% di accuratezza sia in modalità di condizionamento naturale che forzata. Si è inoltre riscontrato un aumento dell'accuratezza quando vengono integrati i dati ambientali e fisiologici (fig.6), mostrando come, per la previsione del comfort termico, i soli dati ambientali non sono sufficienti.

Prediction accuracy of the seven participants in the multi-occupancy room (numbers in bold represent the highest classification accuracy).

Subject ID	1	2	3	4	5	6	7
Envir Only	0.634	0.739	0.655	0.185	0.554	0.529	0.483
Human Only	0.373	0.434	0.601	0.696	0.712	0.679	0.634
Envir + Human	0.932	0.964	0.839	0.723	0.744	0.779	0.748

Figura 6 - Valori di accuratezza utilizzando dati ambientali, fisiologici o integrati.

L'ultimo articolo che verrà analizzato e che è stato maggiormente tenuto in considerazione per lo sviluppo della tesi è "HEAT - Human Embodied Autonomous Thermostat" [5]. Questo studio considera i soggetti che occupano l'ambiente indoor come l'incarnazione di *termostati* intelligenti e connessi, i cui dati fisiologici (come la temperatura cutanea del viso) possono essere utilizzati per controllare le operazioni del sistema HVAC, riuscendo così ad ottenere una migliore soddisfazione termica generale e una riduzione del consumo di energia. Anche in questo caso quindi si cerca di integrare i dati fisiologici ai dati ambientali, per ottenere una previsione accurata della percezione del comfort degli utenti, ovvero del Thermal Sensation Vote (TSV).

I test sono stati effettuati in un ufficio senza finestre e dotato di un termostato che variava la temperatura in un range da 22 °C a 28°C. I soggetti in esame sono 10 studenti universitari di età compresa tra i 22 e i 27 anni. Sono stati effettuati test di raffreddamento e riscaldamento della durata di 60 minuti (fig.7).

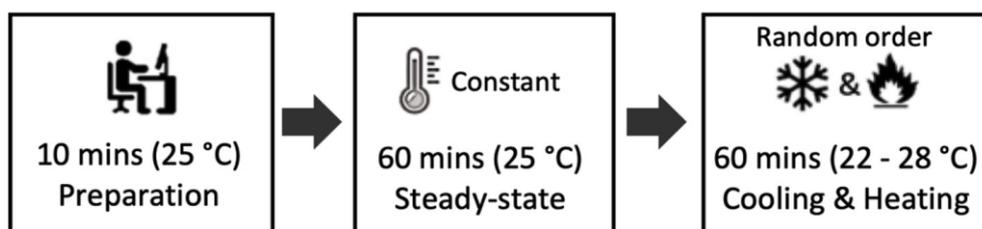


Figura 7 - Protocollo dell'esperimento

Durante le prove sono stati rilevati parametri ambientali (temperatura della stanza, T_{ROOM}) e parametri fisiologici (temperatura della pelle del viso, $T_{SKIN-VISO}$). Quest'ultimi sono stati acquisiti tramite una termocamera *FLIR Lepton*. Inoltre, ogni 3 minuti veniva chiesto ai partecipanti di compilare il questionario per la raccolta dei valori del TSV.

Per l'elaborazione dei dati si sono sviluppate due tipologie di modelli distinti: *modelli di comfort termico personale* e *modelli predittivi fisiologici*. I modelli di comfort termico personale sono stati sviluppati attraverso la funzione di Logistic Regression. Utilizzando tale funzione è possibile ottenere delle previsioni del TSV al variare della temperatura della pelle del volto. Sono state utilizzate diverse zone della faccia come la fronte, la bocca, il naso, le orecchie e le guance. Quest'ultime hanno permesso di ottenere livelli di accuratezza più elevati e dunque per l'analisi è stata maggiormente presa in considerazione la temperatura delle guance (T_{CHEEK}). Il Thermal Sensation Vote predetto dalla Logistic Regression è stato quindi suddiviso in 3 diverse classi: *freddo (-1)* che racchiude i valori di TSV da -1 a -3, *comfort (0)* e *caldo (1)* che racchiude i valori tra +1 e +3. Si sono così ottenuti

grafici in cui vi si ritrovano 3 diverse curve, rappresentanti la probabilità che il soggetto si trovi in discomfort legato al freddo (-1), in comfort (0) o in discomfort legato al caldo (1), al variare della temperatura delle guance (T_{CHEEK}), fig.8.

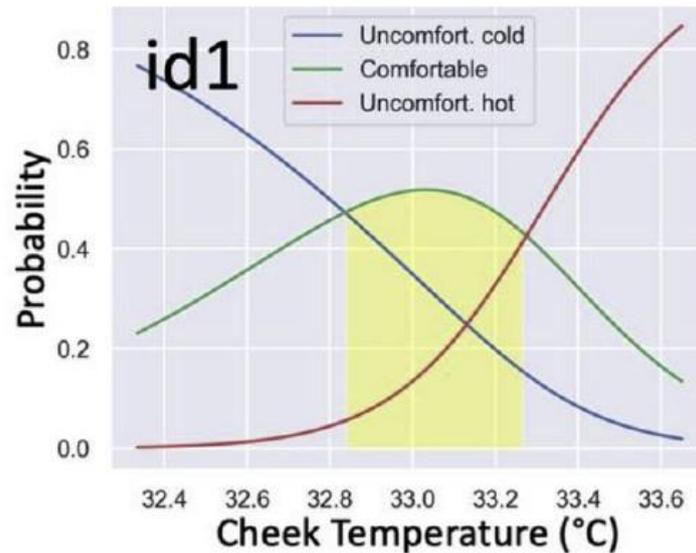


Figura 8 - Modello di comfort termico personale al variare della T_{CHEEK}

I modelli predittivi fisiologici invece sono stati sviluppati attraverso la *Linear mixed model (LMM)*. La LMM, a differenza della Logistic Regression, non considera solo le variazioni delle variabili di interesse di input (effetti fissi), ma anche gli effetti casuali, cioè le variazioni risultanti da campioni casuali della popolazione. In questo caso si prevede la temperatura della pelle del volto (in particolare la T_{CHEEK}) al variare della temperatura della stanza (T_{ROOM}) e ciò permette di poter valutare l'effetto di un nuovo possibile Set-point della T_{ROOM} prima di implementarlo (fig.9).

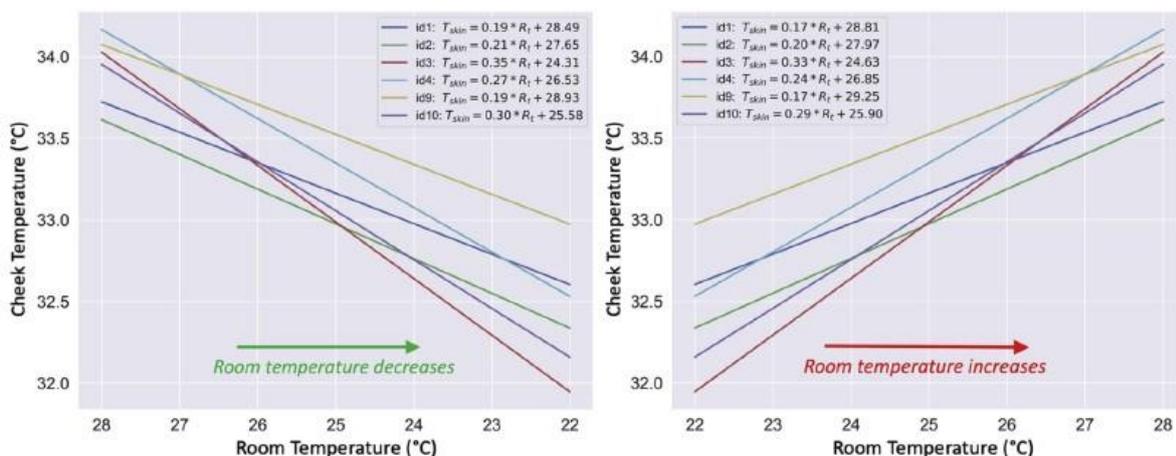


Figura 9 - Modelli predittivi fisiologici nel caso di T_{ROOM} crescente e decrescente.

Una volta ottenuti i modelli si attuano 3 diverse strategie di controllo ed ottimizzazione del comfort termico: *Strategy 1*, *Strategy 2*, *Strategy 3*.

Nella *Strategy 1* viene impostato un Set-point di temperatura della stanza con l'obiettivo di massimizzare la percentuale di utenti in condizione di comfort. Tale Set-point viene definito a partire dalle zone di T_{ROOM} corrispondenti al comfort per ogni soggetto, individuando la zona di temperatura comune a tutti gli utenti, fig.10. Nel caso in figura il Set-point ottimale si trova nel range di T_{ROOM} compreso tra 25,3 °C e 25,4 °C

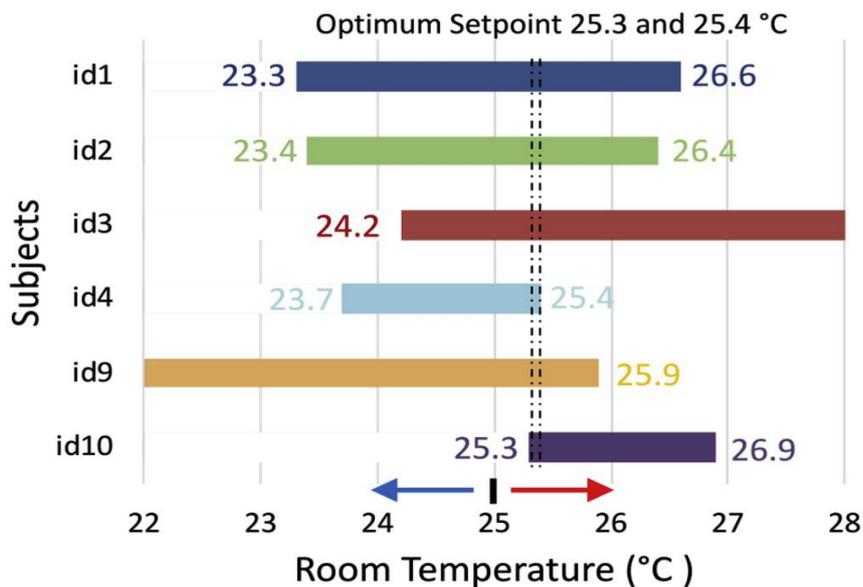


Figura 10 - Set-point ottimale utilizzando *Strategy 1*

Utilizzando invece la *Strategy 2* si cerca di massimizzare anche la percentuale di comfort medio dei soggetti. Tramite l'utilizzo dei modelli di comfort personale si sono ottenute curve di probabilità di comfort per tutti gli utenti. In questa strategia si va quindi ad estrapolare la *curva complessiva di probabilità di comfort* (al variare di T_{ROOM}) operando una “media” dalle curve di probabilità dei singoli soggetti (fig.11). La curva complessiva avrà dunque un punto di massimo e la temperatura della stanza corrispondente sarà dunque il Set-point ottimale, 25,4°C.

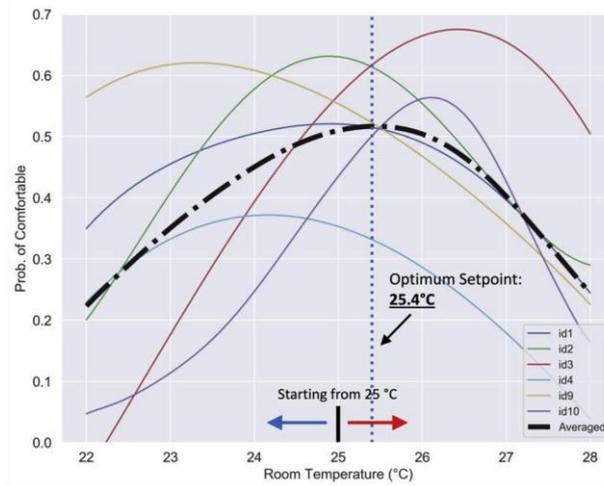


Figura 11 - Set-point ottimale utilizzando Strategy 2

L'ultima strategia di ottimizzazione è la Strategy 3. La metodologia applicata è la medesima della Strategy 2 ma con il vincolo del consumo energetico. In questo caso verranno dunque valutati due coefficienti: il *comfort score* e l'*energy score*, al variare della T_{ROOM} (fig.12)

The comfort score and energy score at different setpoints for a multi-occupancy environment with subjects 1, 2, and 3.

R_t^*	24.2	24.3	24.4	24.5	24.6	24.7	24.8
$\sum_i Prob_i(R_t^*) / n$	0.51	0.52	0.53	0.54	0.55	0.56	0.56
$\sum_i Comfort_score(1)$	0.88	0.90	0.92	0.93	0.94	0.96	0.97
$R_{base} - R_t^*$	-2.2	-2.3	-2.4	-2.5	-2.6	-2.7	-2.8
$Energy_score(2)$	0.37	0.38	0.40	0.42	0.43	0.45	0.47
R_t^*	24.9	25.0	25.1	25.2	25.3	25.4	25.5
$\sum_i Prob_i(R_t^*) / n$	0.57	0.57	0.57	0.58	0.58	0.58	0.58
$\sum_i Comfort_score(1)$	0.98	0.98	0.98	0.99	0.99	1.00	1.00
$R_{base} - R_t^*$	-2.9	-3	-3.1	-3.2	-3.3	-3.4	-3.5
$Energy_score(2)$	0.48	0.50	0.52	0.53	0.55	0.57	0.58

Figura 12 - Energy score e Comfort score al variare di T_{ROOM}

Il maggior contributo del framework HEAT sviluppato in questo studio consiste dunque nell'*ottimizzazione* e nel *rilevamento* del comfort termico a partire dall'utilizzo di parametri fisiologici quale la temperatura della pelle del viso, in particolare delle guance (T_{CHEEK}). Esse infatti, se utilizzate come parametro di input nei modelli di comfort termico personale, permettono di ottenere una previsione del comfort con buoni valori di accuratezza, fig.13.

Prediction accuracy of personal comfort models.

Id	1	2	3	4	5	6	7	8	9	10
Accuracy	0.70	0.78	0.70	0.81	0.41	0.78	0.51	0.58	0.66	0.81

Figura 13 - Accuratezza della previsione del comfort utilizzando la T_{CHEEK}

3. MATERIALI E METODI

3.1. Test preliminare

La prima parte dello studio è stata effettuata su dati precedentemente acquisiti dai ricercatori del dipartimento DIISM dell'Università Politecnica delle Marche. I test sono stati effettuati in una test-room a Bilbao, in Spagna, nell'anno 2019, su 5 soggetti distinti. La durata delle prove è stata di 90 minuti e sono state svolte singolarmente dagli utenti, quindi in un ambiente ad occupazione singola. La stanza è stata inizialmente riscaldata e successivamente raffreddata, con una temperatura variabile tra i 15°C e i 26 °C, come mostrato in fig.14.

Durante i test sono stati rilevati dati ambientali quali la temperatura (t_a), l'umidità (RH) e la velocità dell'aria (v_a). Per tutta la durata della prova è stato richiesto ai partecipanti di compilare un questionario in cui esprimere la propria sensazione termica attraverso il Thermal Sensation Vote (TSV).

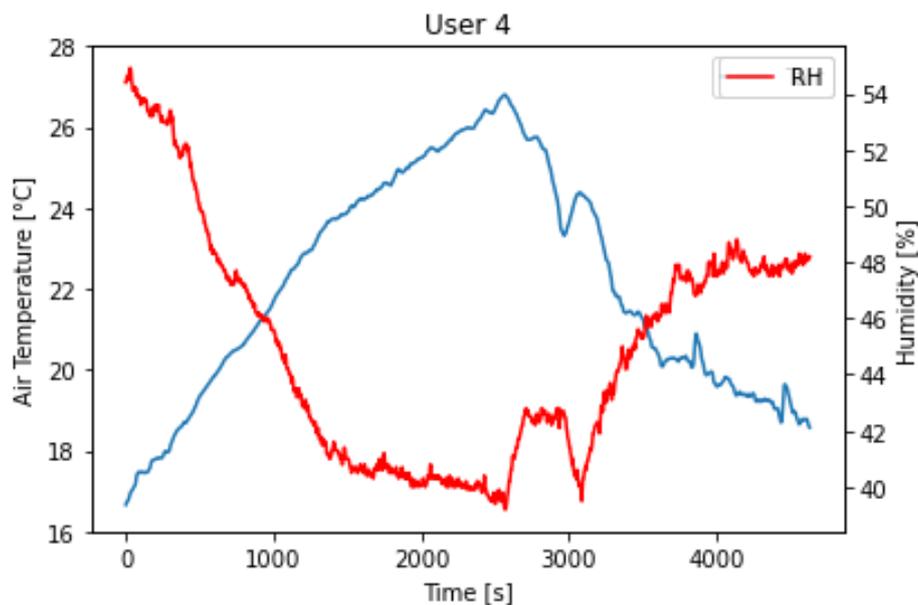


Figura 14 - Variazioni di t_a e RH durante il test

Si sono così potuti osservare gli andamenti del TSV (e dunque della percezione termica dei partecipanti), al variare di t_a , RH e v_a . Nella figura 15 si può osservare il TSV al variare dei parametri ambientali per l'intera durata del test nel caso del soggetto 4.

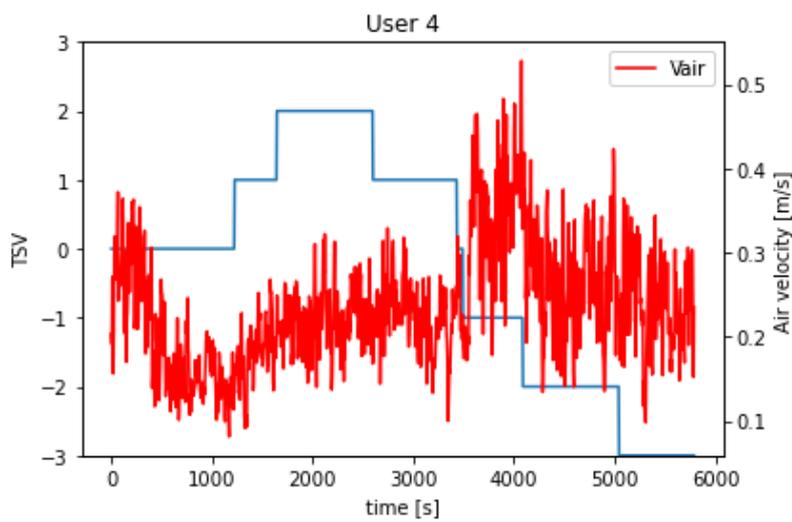
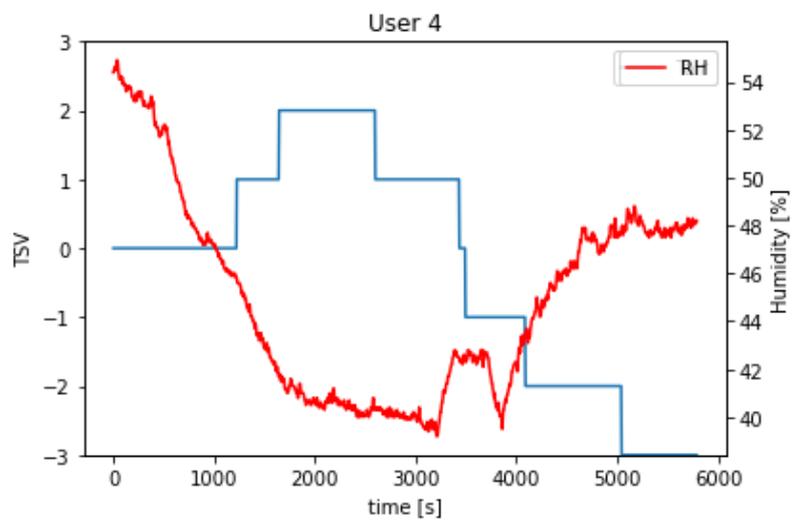
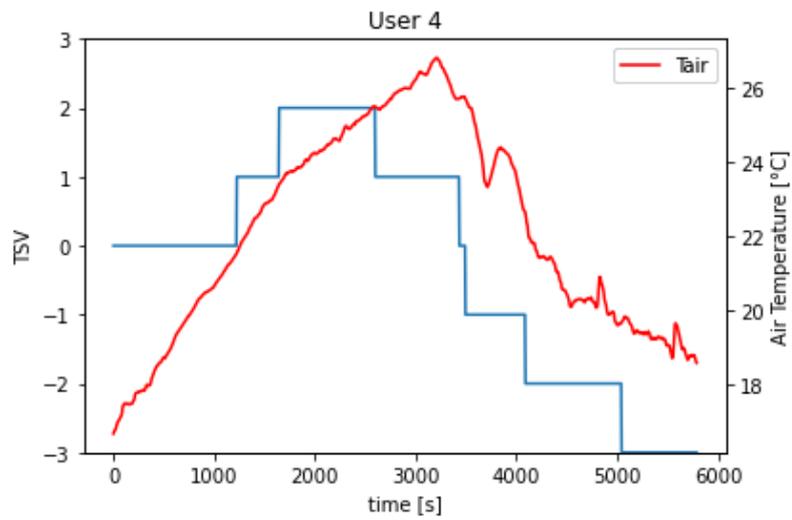


Figura 15 - TSV al variare dei parametri ambientali.

Questa prima parte di studio è servita quindi per sviluppare e convalidare l'algoritmo di Logistic Regression che è stato utilizzato per ottenere previsioni del TSV al variare dei parametri ambientali. Esso è stato allenato singolarmente su temperatura, umidità e velocità dell'aria e sui valori di TSV raccolti nei test. Tramite queste previsioni si sono poi ottenute curve di probabilità di comfort simili a quelle già incontrate per i modelli di comfort termico personali [5].

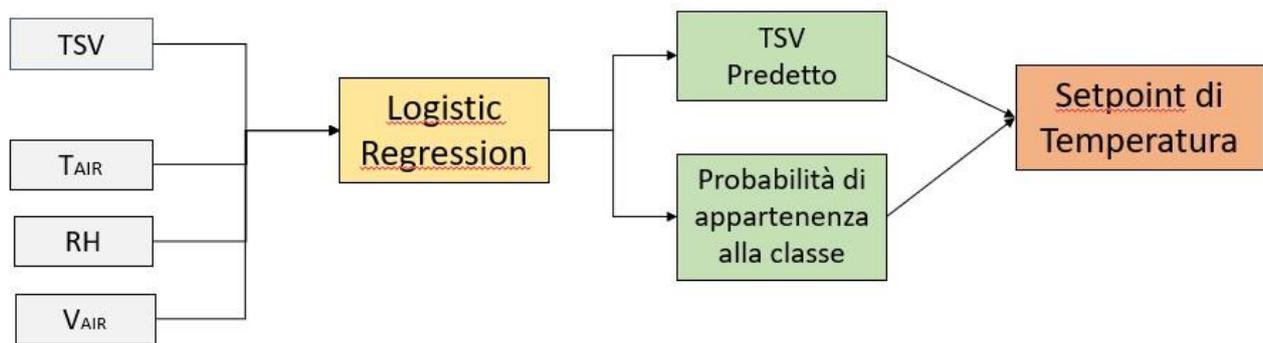


Figura 16 – Procedura utilizzata per allenare l'algoritmo

3.2. Test con termocamera

3.2.1. Descrizione del progetto

Dopo aver svolto uno studio iniziale nel test preliminare, nel quale è stata valutata la percezione del comfort al variare dei soli parametri ambientali, il progetto è proseguito indagando la possibilità di ottenere previsioni del comfort a partire da parametri fisiologici, legati dunque al fattore umano. I parametri che sono stati valutati sono la temperatura della fronte (T_{FRONTE}) e la temperatura delle guance (T_{GUANCE}) e per la loro misurazione è stata utilizzata una termocamera.

L'obiettivo è stato dunque quello di analizzare le relazioni tra percezione termica dell'utente, espressa dal TSV, e tali parametri fisiologici, per ottenere, tramite l'algoritmo di Logistic Regression, una previsione accurata del TSV. Inoltre, anche in questo caso si sono ottenute delle curve di probabilità di comfort al variare di T_{FRONTE} e T_{GUANCE} . Una volta ottenute le previsioni e le curve probabilistiche si procederà poi con la valutazione delle performance dell'algoritmo per entrambe le variabili di input e si sceglierà dunque il parametro fisiologico che garantisce una migliore previsione del comfort termico.

3.2.2. Set-up sperimentale

I test si sono svolti in un ufficio (fig.17) del dipartimento DIISM dell'Università Politecnica delle Marche (Ancona), nel mese di Novembre 2020. I partecipanti alle prove sono stati 6 (5 uomini e 1 donna) di età compresa tra i 24 e 26 anni (tabella 1). Le prove sono state svolte singolarmente dagli utenti e la durata di ogni test è stata di 30 minuti, con una fase iniziale di raffreddamento della stanza (15 minuti), seguita da un riscaldamento per i restanti 15 minuti. Le finestre sono rimaste chiuse per tutta la durata della prova e il raffreddamento e riscaldamento sono stati effettuati rispettivamente con un condizionatore e una stufa elettrica. Il soggetto si trovava seduto (dunque livello di attività stazionario) con una termocamera posta ad 1 metro di distanza ed un sensore di temperatura e umidità posizionato sul tavolo (figura 18). Sono stati così raccolti sia parametri ambientali (t_a e RH) che fisiologici (temperatura del volto tramite la termocamera). La frequenza di campionamento dei dati è stata di 30 secondi e si sono così ottenuti 60 campioni per ogni prova. Con la stessa frequenza è stato inoltre richiesto al soggetto di compilare un questionario relativo al TSV, grazie al quale è stato possibile valutare la sensazione termica.

La temperatura della stanza (t_a) ha subito variazioni in un range contenuto tra 21 °C e 26 °C, mentre per l'umidità (RH) si è osservata un'oscillazione dal 40 % al 50 % (fig.19).

Soggetto	Età	Sesso
1	25	maschio
2	24	maschio
3	24	femmina
4	25	maschio
5	25	maschio
6	26	maschio

Tabella 1 – Informazioni dei partecipanti

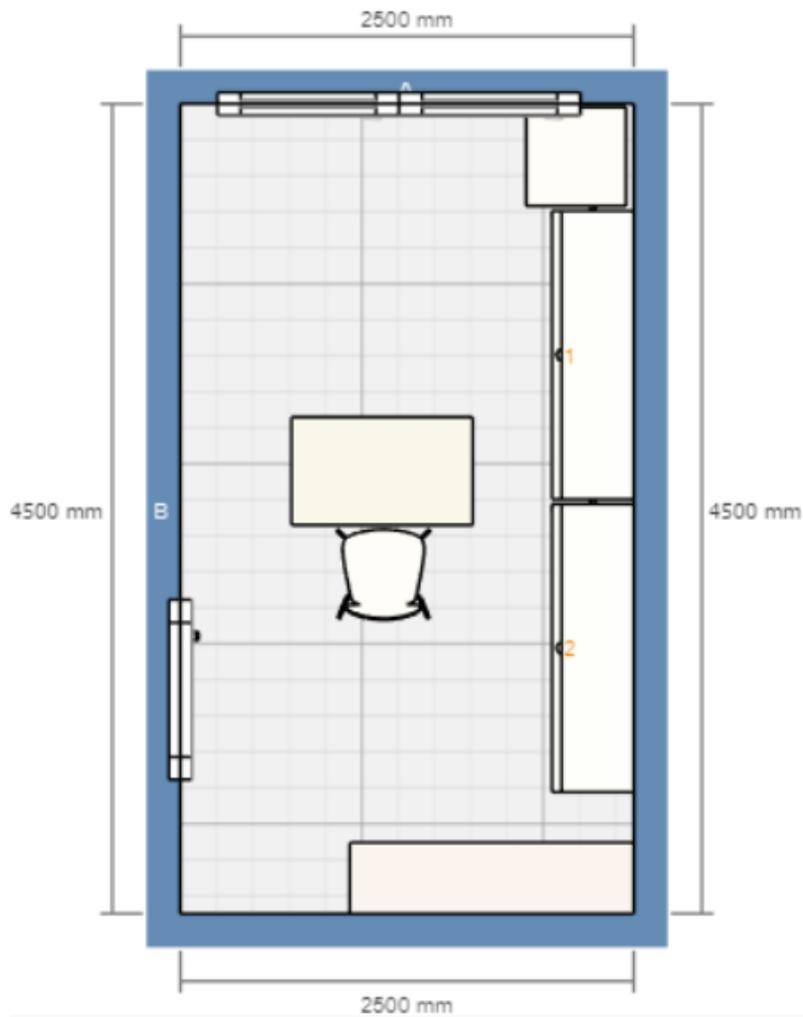


Figura 17 - Piantina ufficio



Figura 18 - Set-up sperimentale

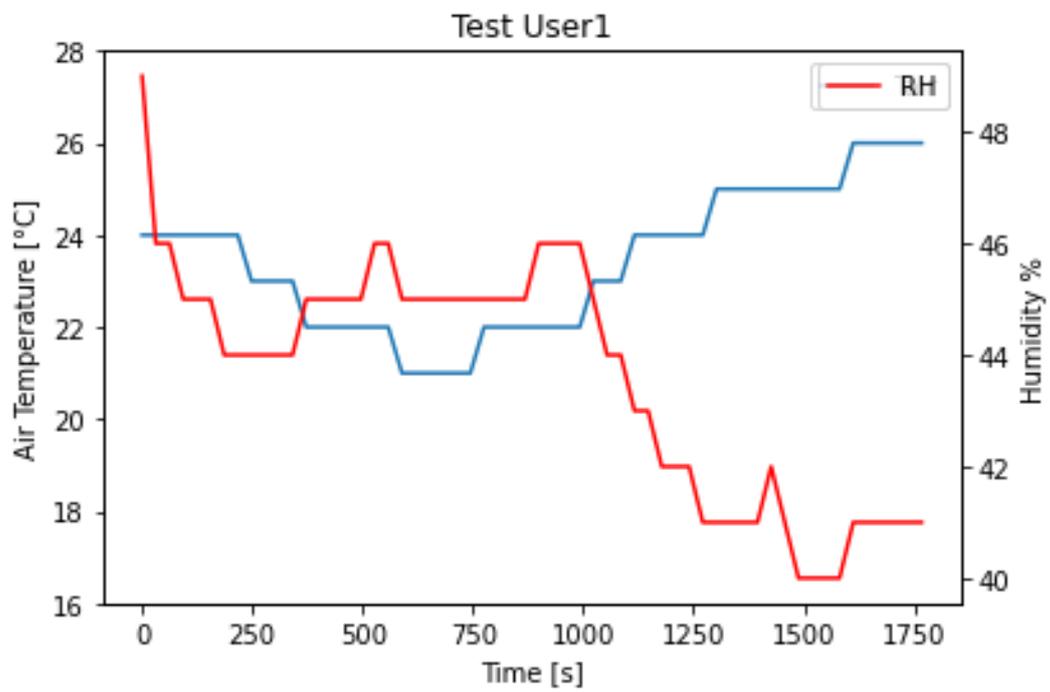


Figura 19 - Variazioni di temperatura e umidità dell'aria durante i test

3.2.3. Sensori utilizzati e acquisizione dati

Per la rilevazione dei parametri durante le prove sono stati utilizzati due strumenti differenti: un sensore *DHT11* ed una termocamera *Flir ThermoCAM S40*.

Il sensore DHT11(fig.20) è stato utilizzato per la raccolta dei dati ambientali, quali la t_a e l'RH. Tale strumento presenta un range di misura della temperatura che si attesta tra gli 0 °C e i 50 °C e dell'umidità compresa tra il 20 % e il 90 %. Per quanto riguarda la temperatura, esso possiede un'accuratezza di misura di ± 1 °C, mentre relativamente all'umidità si ha un'accuratezza di $\pm 1\%$. Il DHT11 è stato collegato ad un computer (sul tavolo a sinistra nella figura 18) e, grazie all'utilizzo di un algoritmo Python, si sono potuti raccogliere i dati di t_a e RH all'interno di un file CSV.

L'altro strumento di misura utilizzato è la termocamera *Flir ThermoCAM S40* (fig.21) con la quale si sono ottenute delle immagini termografiche del volto dei vari soggetti, grazie alle quali è stato possibile estrapolare i valori di T_{FRONTE} e T_{GUANCE} , ovvero dei parametri fisiologici. Il range di misura della temperatura è tra -40 °C e 1500°C con un'accuratezza di misura pari a ± 2 °C. Essa è stata collegata ad un pc e le immagini termografiche sono state acquisite utilizzando il software della termocamera stessa, il *ThermaCAM researcher Professional 2.10*.

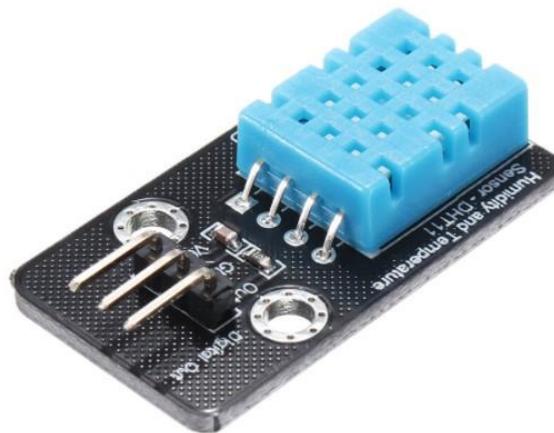


Figura 20 - Sensore DHT11 di temperatura e umidità dell'aria



Figura 21 - Flir ThermoCAM S40

Per quanto riguarda il TSV, esso è stato ottenuto dalla compilazione di un questionario da parte dei soggetti, chiamati a votare da una scala da -3 (freddo) a +3 (caldo). In figura 22 viene mostrato l'andamento del voto di un soggetto (user 1) per tutta la durata della prova.

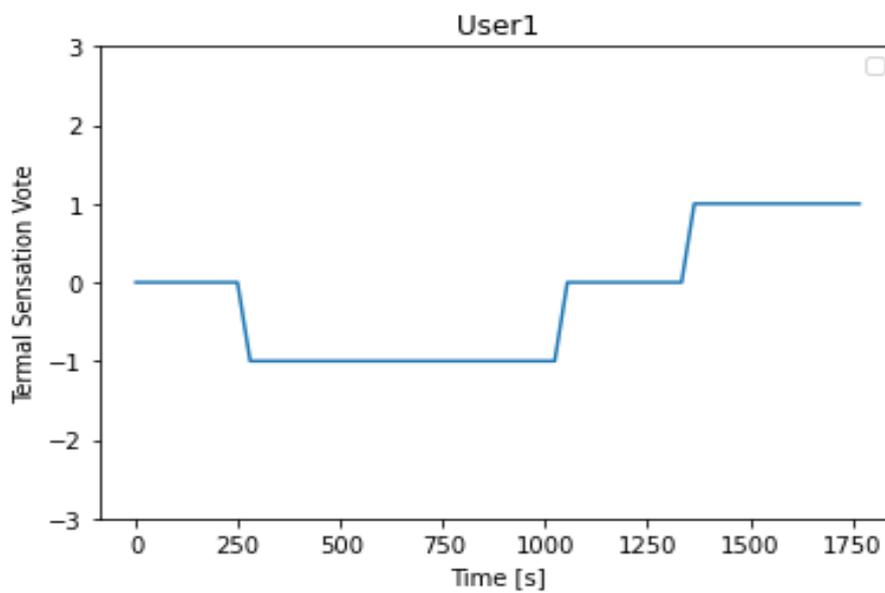


Figura 22 - Variazione del Thermal Sensation Vote durante la prova

Per il processamento delle immagini termografiche si è sviluppato un algoritmo Python tramite il quale è stato possibile estrapolare i parametri fisiologici di interesse, ovvero la temperatura della fronte (T_{FRONTE}) e la temperatura delle guance (T_{GUANCE}).

Verranno ora illustrati i passaggi di cui si compone l'algoritmo, che sono schematicamente racchiusi nella figura 25.

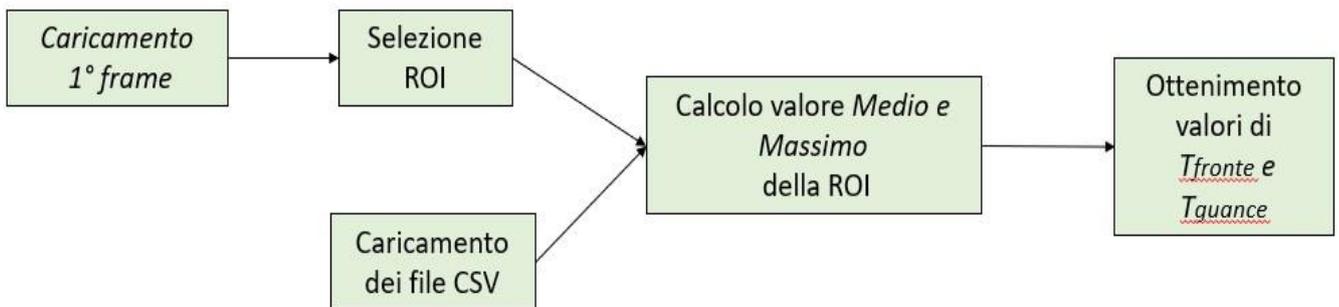


Figura 25 - Step di funzionamento dell'algoritmo per l'analisi delle immagini termografiche.

Si parte con il caricamento dell'immagine termografica in formato JPEG. Ciò è possibile grazie all'utilizzo della libreria *OpenCV*. Per la precisione viene caricato il primo frame (di 60) che verrà utilizzato per lo studio di tutti gli altri. Si ottiene così l'immagine già mostrata in figura 23.

```
img_raw = cv2.imread('immagine.jpg')
```

Una volta caricata l'immagine, sempre grazie all'utilizzo della libreria *OpenCV*, si procede alla selezione della Regione di Interesse (ROI). Tale funzione permette di selezionare manualmente l'area scelta sull'immagine termografica precedentemente caricata. Nel caso di studio è stata selezionata una volta la regione della fronte (per l'ottenimento della T_{FRONTE}) e una volta la regione della guancia (per la T_{GUANCE}), fig.26.

```
roi = cv2.selectROI(img_raw)
print(roi)
roi_cropped = img_raw[int(roi[1]):int(roi[1]+roi[3]), int(roi[0]):int(roi[0]+roi[2])]
```


Si va quindi a calcolare il valore medio e il valore massimo della matrice selezionata per tutti i file CSV (cioè per tutte e 60 le immagini termografiche) e si ottengono quindi i valori di T_{FRONTE} e T_{GUANCE} . Per l'analisi successiva dei dati si è poi deciso di utilizzare i valori massimi dei parametri fisiologici in esame. Questa scelta deriva dalla maggior precisione del valore massimo rispetto al medio infatti, andando a selezionare manualmente la ROI sul primo frame, se la posizione del volto del soggetto nei frame successivi varia anche in maniera sensibile, all'interno dell'area selezionata potrebbe essere inclusa anche una porzione di sfondo (ad esempio il muro) che, avendo una temperatura molto più bassa del volto, influisce sul calcolo del valore medio. Utilizzando il valore massimo, invece, si ha sempre la certezza di rilevare la temperatura di un punto della ROI nel soggetto.

```

massimo=[]
media=[]
for i in range(len(area)):
    massimo.append(np.max(area[i]))
    massimo[i]=np.max(massimo[i][0])
for i in range(len(area)):
    media.append(np.mean(area[i]))
    media[i]=np.round(np.mean(media[i][0]),1)

```

I valori medi e massimi della T_{FRONTE} e della T_{GUANCE} per tutta la durata del test si possono osservare nei grafici di figura 28 e 29.

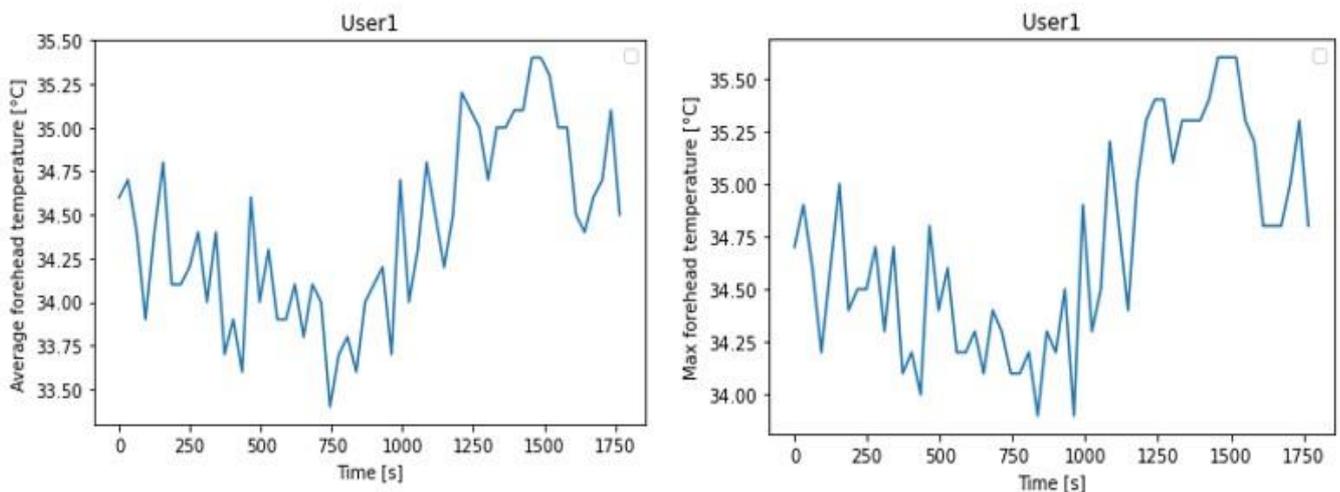


Figura 28 - Valori medi e massimi della T_{FRONTE} durante la prova

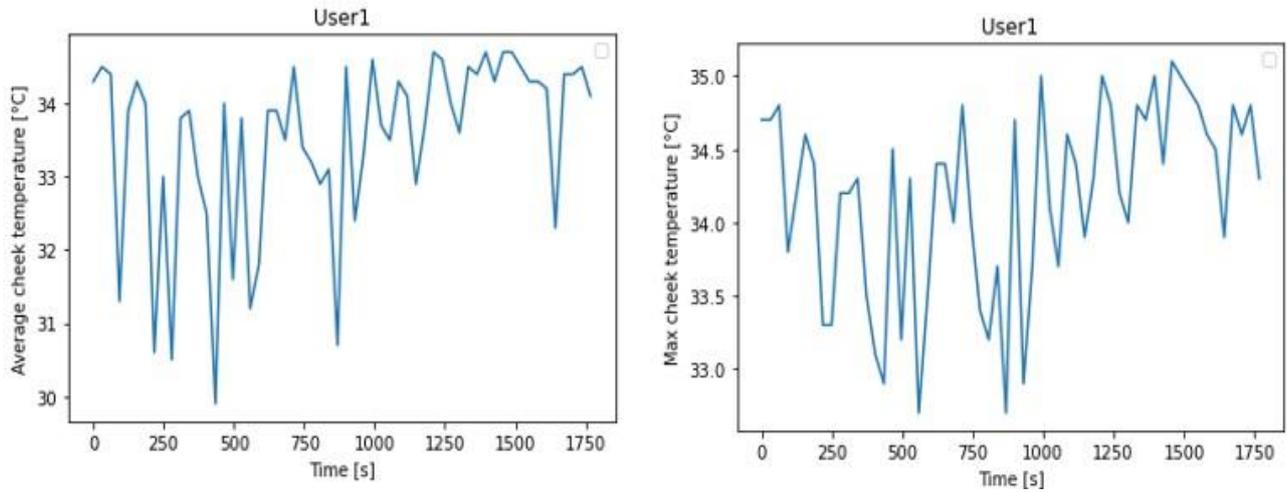


Figura 29 - Valori medi e massimi T_{GUANCE} durante la prova

Una volta ottenuti quindi i dati fisiologici, essi sono raccolti assieme ai dati ambientali (t_a e RH) e ai valori del TSV all'interno di un unico file CSV. Si procede quindi all'analisi dei dati andando a realizzare una previsione del comfort termico (TSV) e ottenendo le curve di probabilità tramite l'algoritmo Python sviluppato durante il test preliminare.

3.4. Analisi dei dati

I dati del test preliminare e quelli raccolti durante il test con la termocamera sono stati elaborati ed analizzati mediante un algoritmo sviluppato utilizzando il linguaggio di programmazione Python. Esso è stato ottenuto facendo ricorso al metodo del *Machine Learning Supervisionato* tramite il quale è stato sviluppato un algoritmo di Logistic Regression, capace poi di predire il TSV degli utenti al variare di parametri ambientali (test preliminare) e parametri fisiologici (test con termocamera).

3.4.1. Machine Learning Supervisionato

Quando si parla di Machine Learning (ML) si fa riferimento a quei metodi mediante i quali si insegna ai computer e ai robot a svolgere azioni ed attività in modo autonomo attraverso programmi di apprendimento automatico. Sostanzialmente, gli algoritmi di ML usano metodi matematico-computazionali per apprendere informazioni direttamente dai dati, senza modelli matematici ed equazioni predeterminate. Gli algoritmi di ML migliorano le loro performance in modo “adattivo” a mano a mano che gli “esempi” da cui apprendere aumentano [7].

Vi sono sostanzialmente due approcci distinti quando si parla di ML che permettono di distinguere l'apprendimento automatico in due sottocategorie a seconda del fatto che si diano all'algoritmo esempi completi da utilizzare come indicazione per eseguire il compito richiesto (apprendimento supervisionato) oppure che lo si lasci lavorare senza alcun “aiuto” (apprendimento non supervisionato), fig.30.

Quando si utilizza l'approccio del ML Supervisionato, come nel contesto di questa tesi, vengono forniti all'algoritmo sia dei set di dati come input sia le informazioni relative ai risultati desiderati (output) così da permettere al sistema di elaborare delle relazioni tra i dati in ingresso e quelli di uscita. Vengono cioè forniti al sistema degli esempi di input e output così che esso possa comprendere il nesso esistente tra di loro [7]. Nell'apprendimento supervisionato il lavoro di risoluzione viene lasciato all'algoritmo. Una volta acquisita la funzione matematica che ha portato a risolvere uno specifico insieme di problemi, sarà quindi possibile riutilizzare la funzione per rispondere a problemi simili [8].

Nel caso invece del ML non supervisionato vengono forniti al sistema solamente dei set di dati senza nessuna informazione sul risultato desiderato. Attraverso questo secondo metodo di apprendimento ci si aspetta infatti di ottenere degli schemi e modelli nascosti, ovvero di identificare all'interno degli input una struttura logica senza che questi siano anticipatamente etichettati [7].

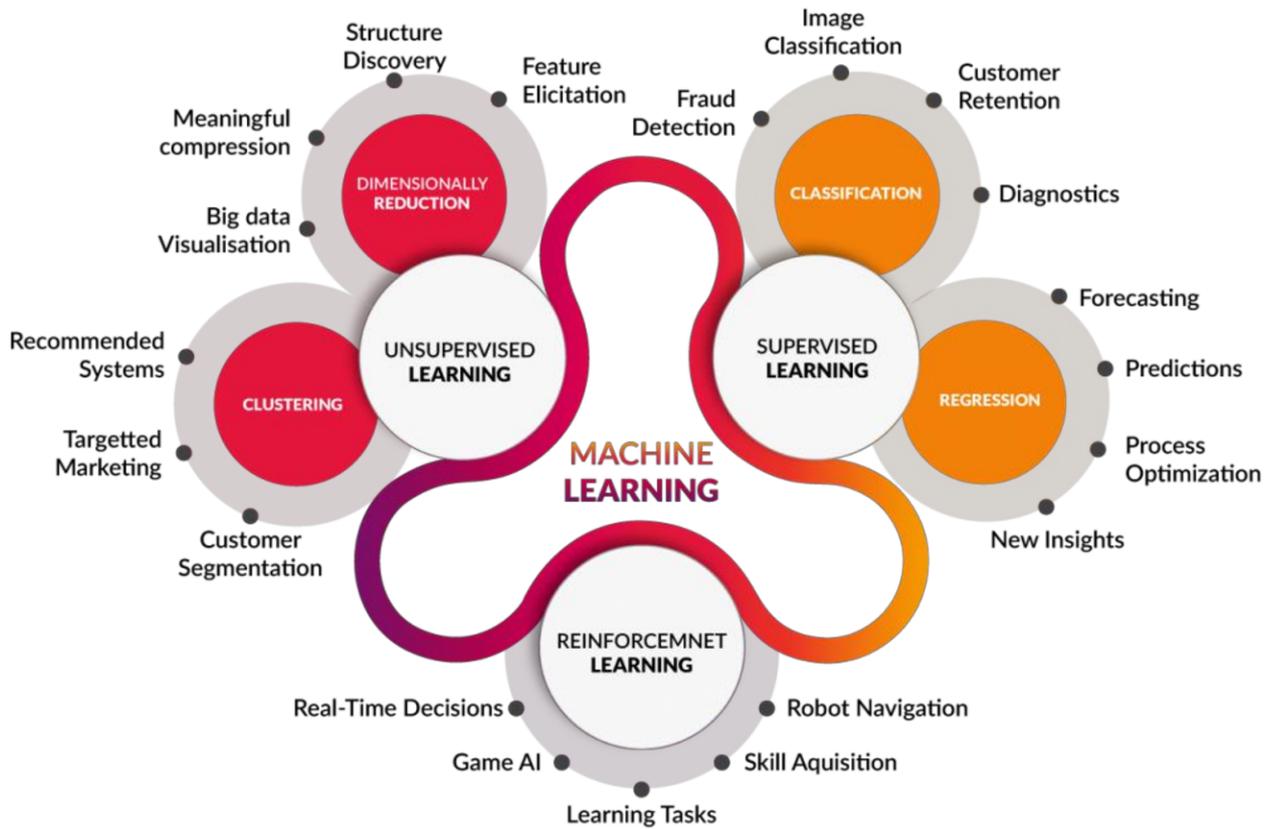


Figura 30 - Machine Learning supervisionato e non supervisionato

Nel caso di studio della tesi è stato quindi utilizzato il metodo del ML Supervisionato nel linguaggio di programmazione Python. Sono stati infatti forniti all’algoritmo esempi di input e output e quest’ultimo è stato poi in grado di “apprendere” la relazione che intercorre tra di loro.

3.4.2. Logistic Regression

L'algoritmo utilizzato per l'analisi dei dati del test preliminare e del test effettuato con la termocamera è stato sviluppato sulla base del modello della Logistic Regression.

Tale modello è uno strumento molto importante nella disciplina del Machine Learning, infatti si può considerare la regressione logistica come un metodo di classificazione rientrante nella famiglia degli algoritmi di apprendimento supervisionato. Essa si avvale dell'utilizzo di un algoritmo per classificare i dati in arrivo in base a dati storici e, a mano a mano che vengono analizzati più dati, l'algoritmo dovrebbe migliorare nel prevedere le classificazioni [9]. Per questo motivo la regressione logistica è infatti diventata particolarmente popolare nella pubblicità online, permettendo ai marketer di prevedere, come percentuale sì o no, la probabilità che utenti specifici di un sito web facciano clic su determinati annunci pubblicitari [9].

Grazie all'utilizzo di metodi statistici, la Logistic Regression permette di generare un risultato che, di fatto, rappresenta una *probabilità* che un dato valore di ingresso appartenga a una determinata classe [9]. Come tutte le analisi di regressione, la regressione logistica è un'analisi predittiva che viene utilizzata per misurare la relazione tra la *variabile dipendente* (ossia ciò che si desidera prevedere) e l'una o più *variabili indipendenti* (le caratteristiche), stimando delle probabilità tramite una funzione logistica [9]. Tale funzione logistica, anche detta funzione sigmoidea, è una curva a forma di S che può prendere qualsiasi numero di valore reale e mapparlo in un valore compreso tra 0 e 1, estremi esclusi [9]. Essa può essere rappresentata tramite la seguente equazione (1):

$$F(x) = \frac{1}{1+e^{-(b_0+b_1*x)}} \quad (1)$$

Dove:

- e : base dei logaritmi naturali (il numero di Eulero o la funzione EXP () nel foglio Excel)
- b_0 e b_1 : coefficienti dei valori di input (numeri reali costanti)
- x : è il valore numerico effettivo che si desidera trasformare.

L'equazione che rappresenta invece la regressione logistica, ipotizzando di avere solamente due coefficienti, può essere descritta come (2):

$$y = \frac{e^{(b_0+b_1*x)}}{1+e^{(b_0+b_1*x)}} \quad (2)$$

Dove:

- X = valore di input;
- b_0 e b_1 = coefficienti dei valori di input (numeri reali costanti);
- y = valore output da predire.

L'equazione della Logistic Regression può essere espressa anche in termini probabilistici. A partire dall'equazione 2, si parte modellando la probabilità che un input (X) appartenga alla classe predefinita ($Y = 1$). Si può scrivere:

$$P(X) = P(Y = 1 | X)$$

Come già spiegato in precedenza, le previsioni vengono trasformate utilizzando la funzione logistica. Pertanto, ipotizzando che Y , ossia il risultato, appartenga alla classe 1, possiamo sostituire alla y dell'equazione (2) la $P(X)$:

$$P(X) = \frac{e^{(b_0+b_1*X)}}{1+e^{(b_0+b_1*X)}} \quad (3)$$

e dunque, essendo e l'opposto del logaritmo, si ottiene:

$$\ln\left(\frac{p(X)}{1-p(X)}\right) = b_0 + b_1 * X \quad (4)$$

Esistono due tipologie di Logistic Regression: la regressione logistica *binomiale* e *multinomiale*.

Nel caso di regressione logistica *binomiale* (fig.31), la probabilità che l'output appartenga ad una classe sarà P , mentre che appartenga all'altra classe $1-P$ (dove P è un numero compreso tra 0 e 1 perché esprime una probabilità) [9]. Questa tipologia di regressione lavora bene in tutti quei casi in cui la variabile che si sta cercando di predire è binaria, cioè può assumere solamente due valori: il valore 1 che rappresenta la classe positiva, o il valore 0 che rappresenta la classe negativa [9].

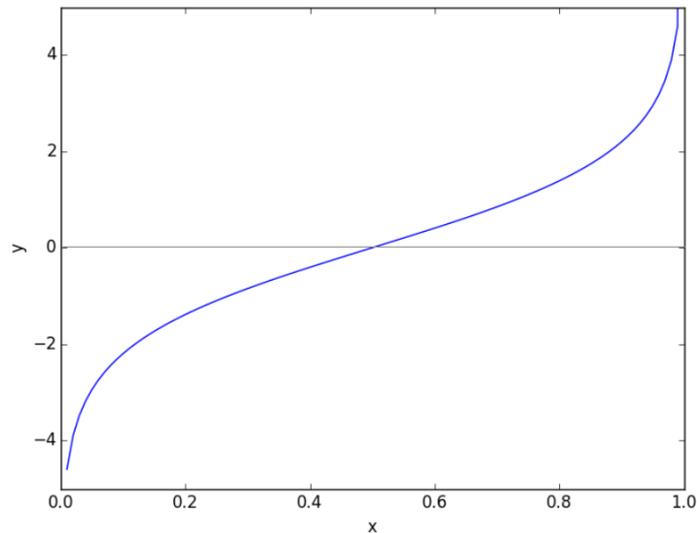


Figura 31- Logistic Regression binomiale

L'altro tipo di regressione logistica invece è la *multinomiale* (fig.32). In questo caso la variabile di risposta può avere tre o più valori possibili ed è quindi in grado di prevedere eventi con classi multiple. Nello studio realizzato in questa tesi si è utilizzata una variabile di input (un parametro ambientale o fisiologico) e una variabile di output (il TSV) che contiene 3 valori categoriali (classi) diversi, che sono: discomfort legato al freddo, discomfort legato al caldo e la condizione neutrale di comfort.

Per lo sviluppo dell'algorithmo utilizzato per l'analisi dei dati è stata quindi utilizzata una funzione di *Logistic Regression multinomiale*.

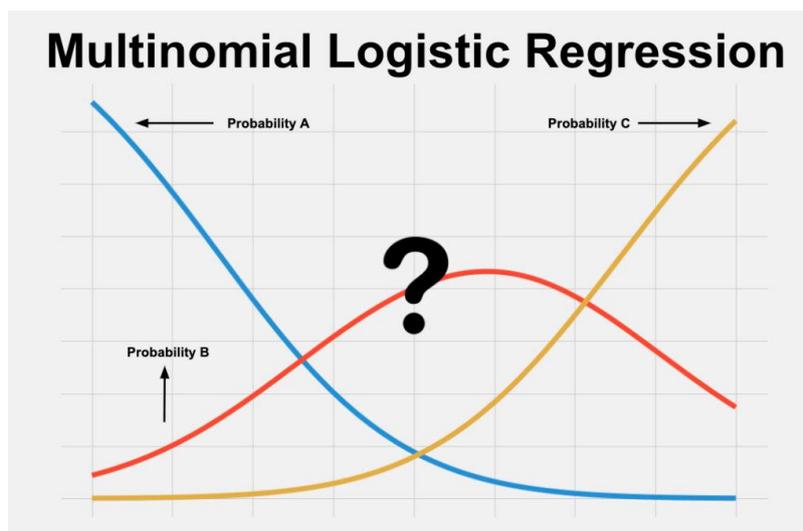


Figura 32 - Logistic Regression multinomiale

3.4.3. Algoritmo per l'analisi dei dati

L'algoritmo utilizzato per l'analisi e l'elaborazione dei dati sia nel caso del test preliminare sia nel test effettuato con la termocamera è stato sviluppato tramite il linguaggio di programmazione Python.

L'obiettivo dello script realizzato è quello di ottenere delle previsioni del Thermal Sensation Vote (TSV) degli utenti al variare di diverse variabili di input (fattori ambientali o fattori fisiologici). Una volta ottenuta la previsione si è proceduto poi all'extrapolazione della probabilità che i vari soggetti appartengano ad una delle 3 classi (freddo, comfort, caldo) nelle quali è stato suddiviso il TSV. Per permettere all'algoritmo di ottenere tali previsioni gli sono stati forniti degli esempi di input e di output tramite i quali è stato poi in grado di 'apprendere' la relazione che intercorre tra di loro (ML supervisionato). Si passa ora all'illustrazione degli step di costruzione dello script.

Tramite l'utilizzo della libreria *pandas* viene caricato il file excel contenente i dati di input e di output.

```
df = pd.read_excel ('file excel')
df.head()
```

Ottenuto il dataframe si passa alla creazione di due variabili: una di input (X) e una di output (Y). La variabile di input può essere uno qualsiasi dei parametri ambientali o fisiologici, mentre la variabile di output è il TSV.

```
X = df['input']
Y = df['output']
```

Come già spiegato, nel caso in esame si è utilizzata una regressione logistica multinomiale poiché la variabile di output (TSV) è stata suddivisa in più classi, per la precisione 3. Si divide infatti il TSV in 3 categorie: la condizione di discomfort legato al freddo (-1), quella di comfort (0) e quella di discomfort legato al caldo (+1). La prima racchiude i valori di TSV tra -1 e -3, la seconda solamente il valore 0, mentre la terza i valori di TSV compresi tra +1 e +3.

```
valori_greater_0 = (Y > 0)
Y[valori_greater_0] = 1
valori_minor_0 = (Y < 0)
Y[valori_minor_0] = -1
Y.unique()
```

I dati di input e di output vengono poi divisi in due parti: un 75 % destinato al *training* e un 25 % destinato al *test*.

```
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size = 0.25)
```

Infatti, essendo questo un algoritmo di apprendimento automatico supervisionato, la Logistic Regression deve essere *allenata* su dati di input e di output, ovvero sui dati di *training*, così da riconoscere una relazione tra essi e riuscire poi a fornire delle previsioni accurate dell'output, che verranno realizzate appunto sui dati di test.

```
lreg = LogisticRegression()
lreg.fit(X_train.values.reshape(-1,1), y_train.values.reshape(-1,1))
y_pred = lreg.predict(X_test.values.reshape(-1,1))
```

La Logistic Regression permette anche di ottenere la *probabilità* che un dato valore di ingresso appartenga a una determinata classe [9]. Nel caso in esame che appartenga ad una delle 3 classi nelle quali è stato suddiviso il TSV (output).

```
y_prob = lreg.predict_proba(X_test.values.reshape(-1,1))
```

Le probabilità di appartenenza alle 3 classi possono essere visualizzate all'interno di un grafico sottoforma di 3 differenti curve che rappresentano le condizioni di freddo, comfort e caldo. Tali probabilità sono ottenute ancora una volta da previsioni dell'output da parte della Logistic Regression sui dati di test.

```
figure3 = plt.figure()
ax1 = figure3.add_subplot(111)
ax1.set_title('Probabilità')
ax1.scatter(X_test, y_prob[:,0], label = 'Cold')
ax1.scatter(X_test, y_prob[:,1], label = 'Comfort')
ax1.scatter(X_test, y_prob[:,2], label = 'Hot')
```

In figura 33 viene mostrato uno schema riassuntivo del funzionamento dell'algoritmo.

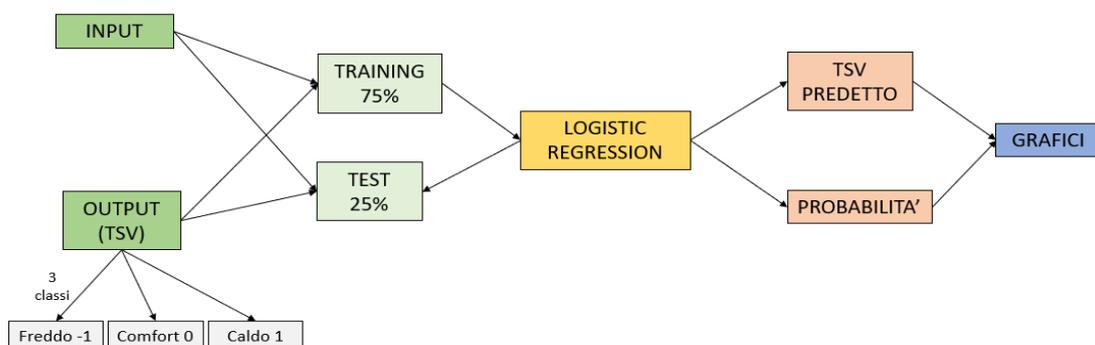


Figura 33 - Principio di funzionamento dell'algoritmo sviluppato

3.4.4. Performance dell'algoritmo

Messo a punto l'algoritmo è il momento di verificarne le prestazioni. La valutazione viene fatta attraverso l'utilizzo di due indici: *Indice di correlazione di Pearson* e indice di *accuratezza*.

L'equazione che esprime l'indice di correlazione lineare di Pearson è:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad (5)$$

Considerate due variabili statistiche X e Y, l'indice di correlazione di Pearson è definito come la loro covarianza divisa per il prodotto delle deviazioni standard [10]. Nel caso di studio, le due variabili sono il TSV reale e il TSV predetto. Tale coefficiente può assumere valori compresi tra -1 e +1, dove il valore 0 indica che tra le due variabili considerate non vi è nessuna correlazione. Il valore +1 invece indica la presenza di massima correlazione lineare in maniera proporzionale mentre quando è pari a -1 indica che la correlazione è inversamente proporzionale [11]. Riassumendo, il coefficiente di Pearson mette in evidenza quanto la dispersione dei punti appartenenti alle due variabili può essere approssimato con una retta, ma non fornisce informazioni sulla sensibilità, cioè sulla sua pendenza. All'interno dello script Python tale indice di correlazione è stato calcolato con il codice:

```
corr, _ = pearsonr(y_test, y_pred)
print('Pearsons correlation: %.3f' % corr)
```

L'altro indice utilizzato è l'*accuratezza*. Esso permette di valutare l'efficacia del modello predittivo e viene espresso come:

$$Accuracy = \frac{TP + TN}{TN + FP + FN + TP} \quad (6)$$

L'accuratezza è definita quindi dal rapporto tra le predizioni esatte (veri positivi e veri negativi) e le predizioni totali effettuate (veri negativi, veri positivi, falsi negativi e falsi positivi) [12]. Sarà dunque un numero racchiuso fra 0 ed 1 che, visto in termini percentuali, esprime l'affidabilità della previsione realizzata dall'algoritmo, con ovviamente una previsione tanto migliore tanto più l'accuratezza tende al valore di 1 (ovvero 100 %).

```
score=lreg.score(x,y)
```

4. RISULTATI

4.1. Risultati del test preliminare

Nella prima parte di studio si è applicato l'algoritmo sopra descritto ai dati raccolti dai ricercatori dell'Univpm a Bilbao. In questo caso si è voluta indagare la percezione termica dei soggetti sottoposti ai test al variare dei parametri ambientali. L'algoritmo di Logistic Regression è stato quindi allenato singolarmente su una variabile ambientale di input alla volta. Si è lavorato dunque su 3 dataset distinti: uno riferito alla temperatura dell'aria (t_a), un altro all'umidità RH e l'ultimo alla velocità dell'aria (v_a). Per ognuno dei 3 casi si sono poi ottenuti dei grafici che mostrano la probabilità che il TSV appartenga ad una delle 3 classi (freddo, comfort, caldo). Si hanno quindi 3 grafici di probabilità per ogni utente, che sono appunto 5. Inoltre, si sono ricavati grafici che descrivono gli andamenti del TSV reale e del TSV predetto dalla Logistic Regression. Quest'ultimi sono utili per verificare anche visivamente la correlazione tra previsione e il valore reale, che verrà poi esplicitata dal calcolo dell'indice di correlazione di Pearson. Per ogni parametro ambientale di input è stato dunque calcolato il coefficiente di Pearson e l'accuratezza, tramite la quale è stato possibile valutare l'efficacia del modello predittivo.

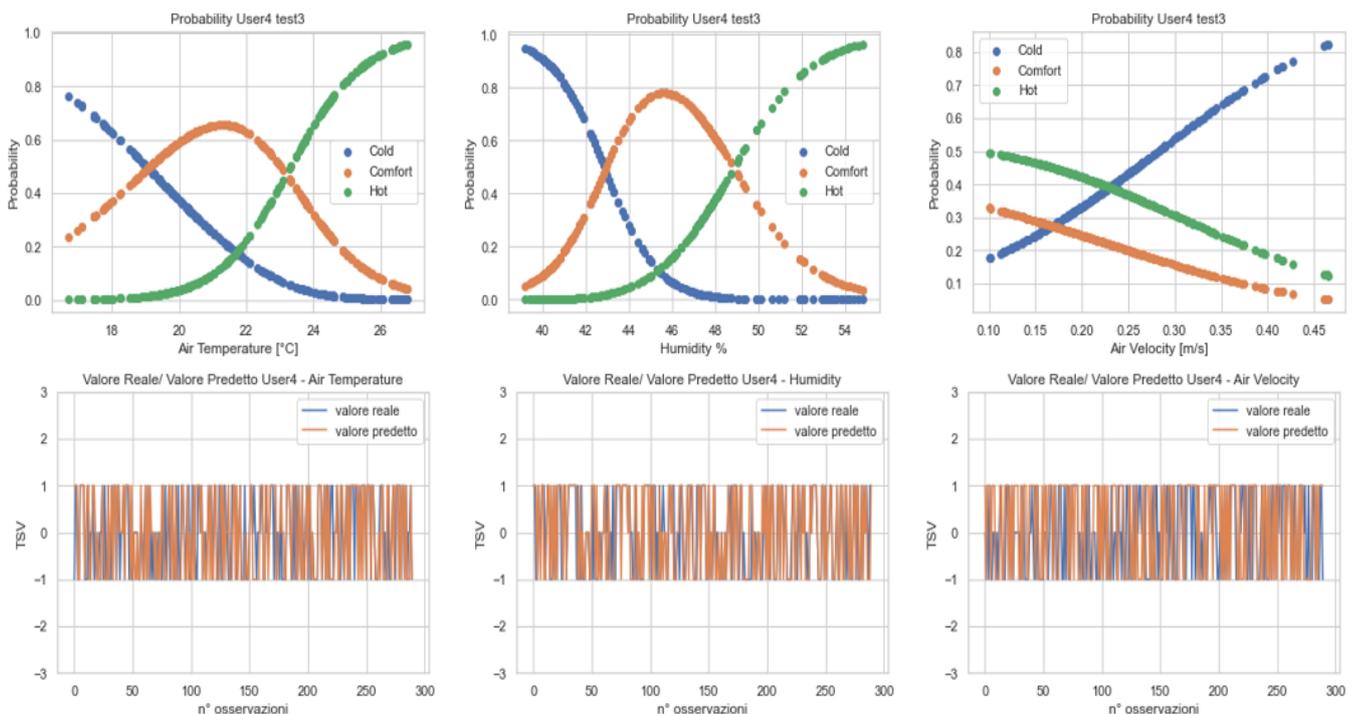


Figura 34 - Grafici di probabilità e di correlazione del soggetto 4

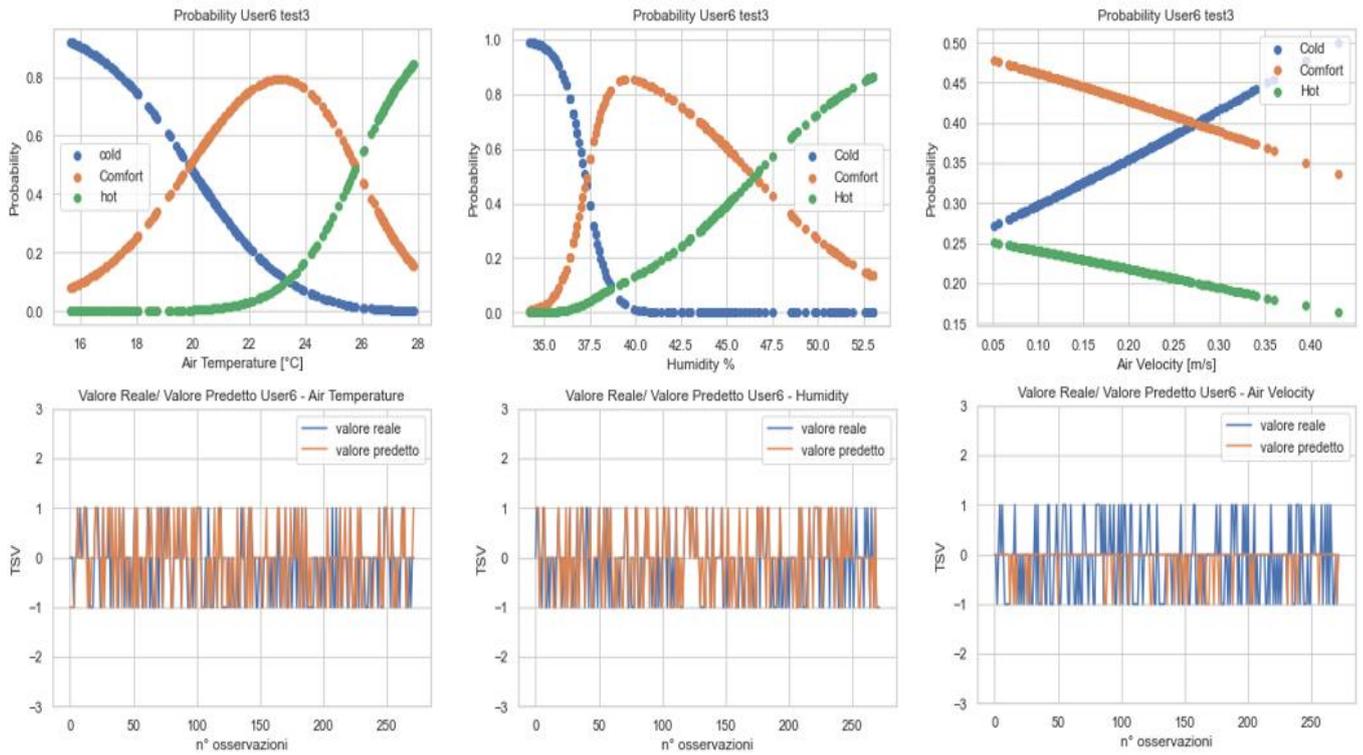


Figura 35 - Grafici di probabilità e di correlazione del soggetto 6

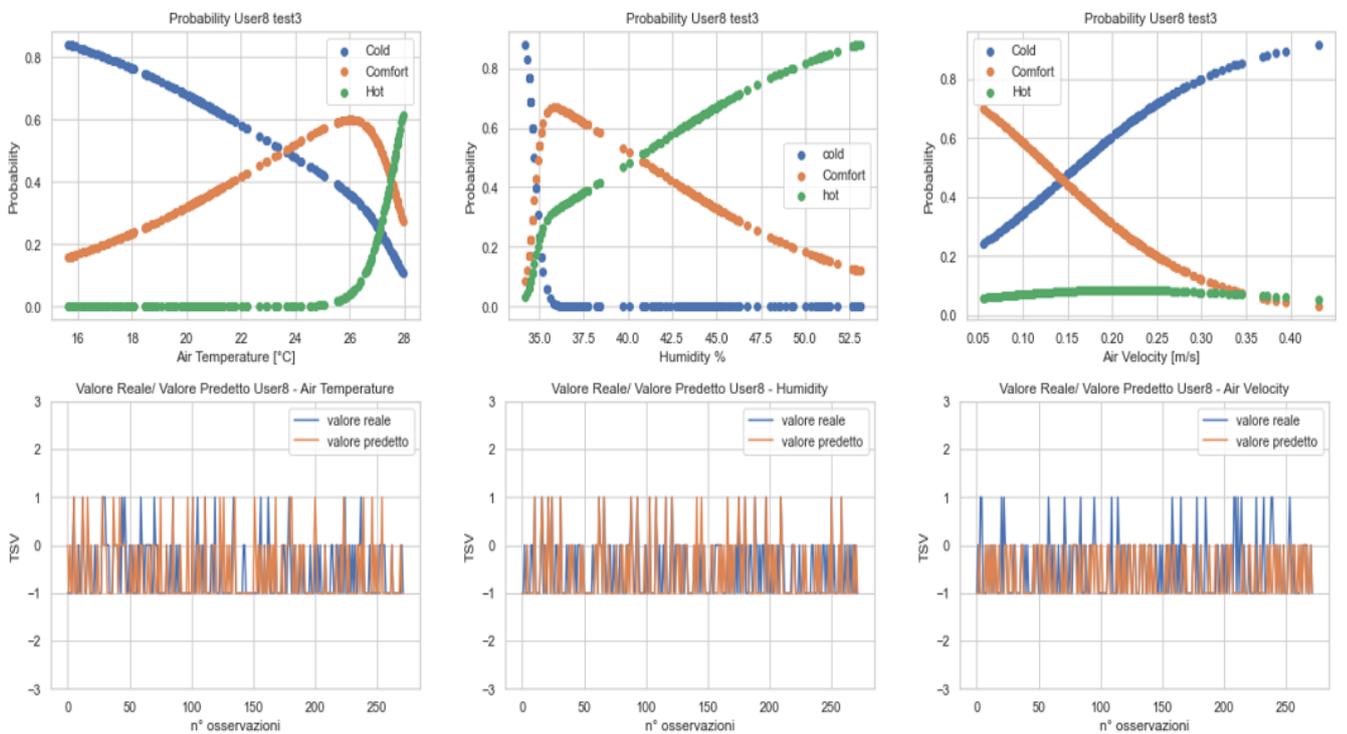


Figura 36 - Grafici di probabilità e di correlazione del soggetto 8

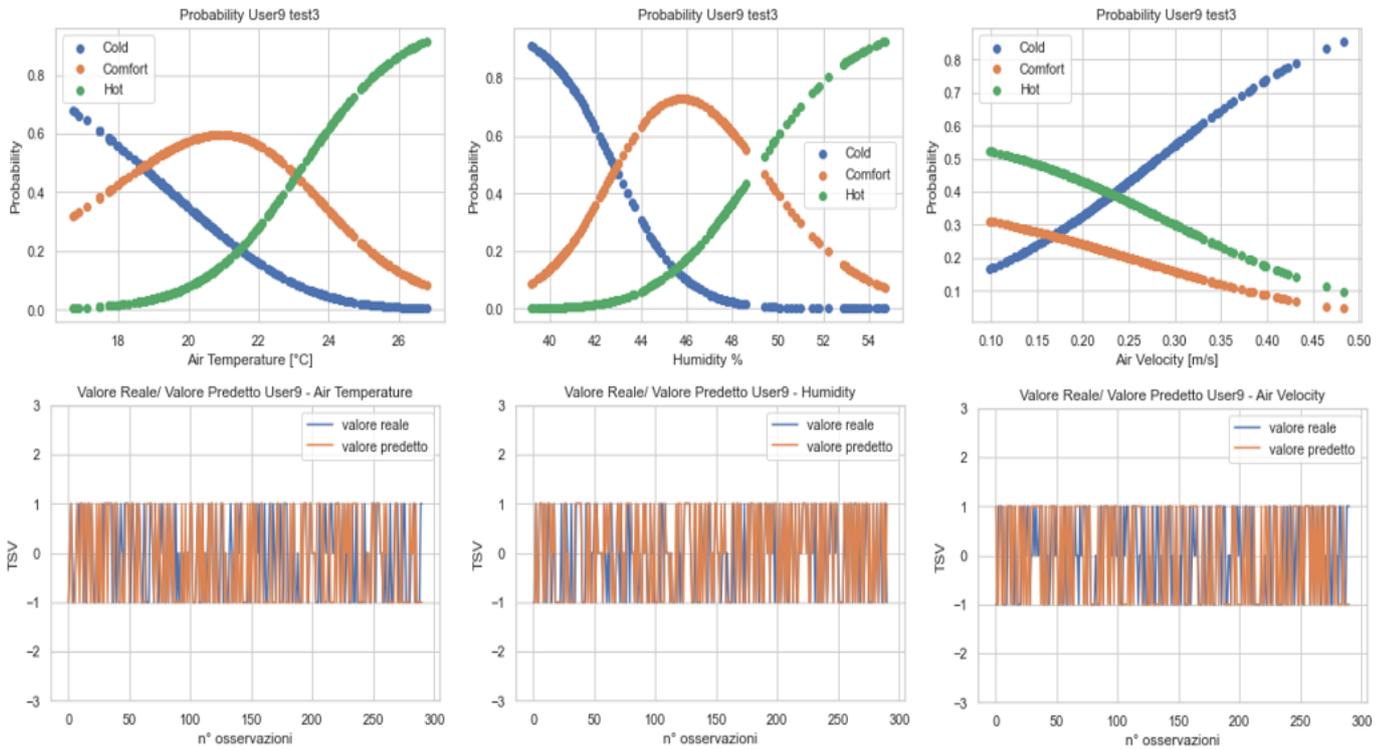


Figura 37 - Grafici di probabilità e di correlazione del soggetto 9

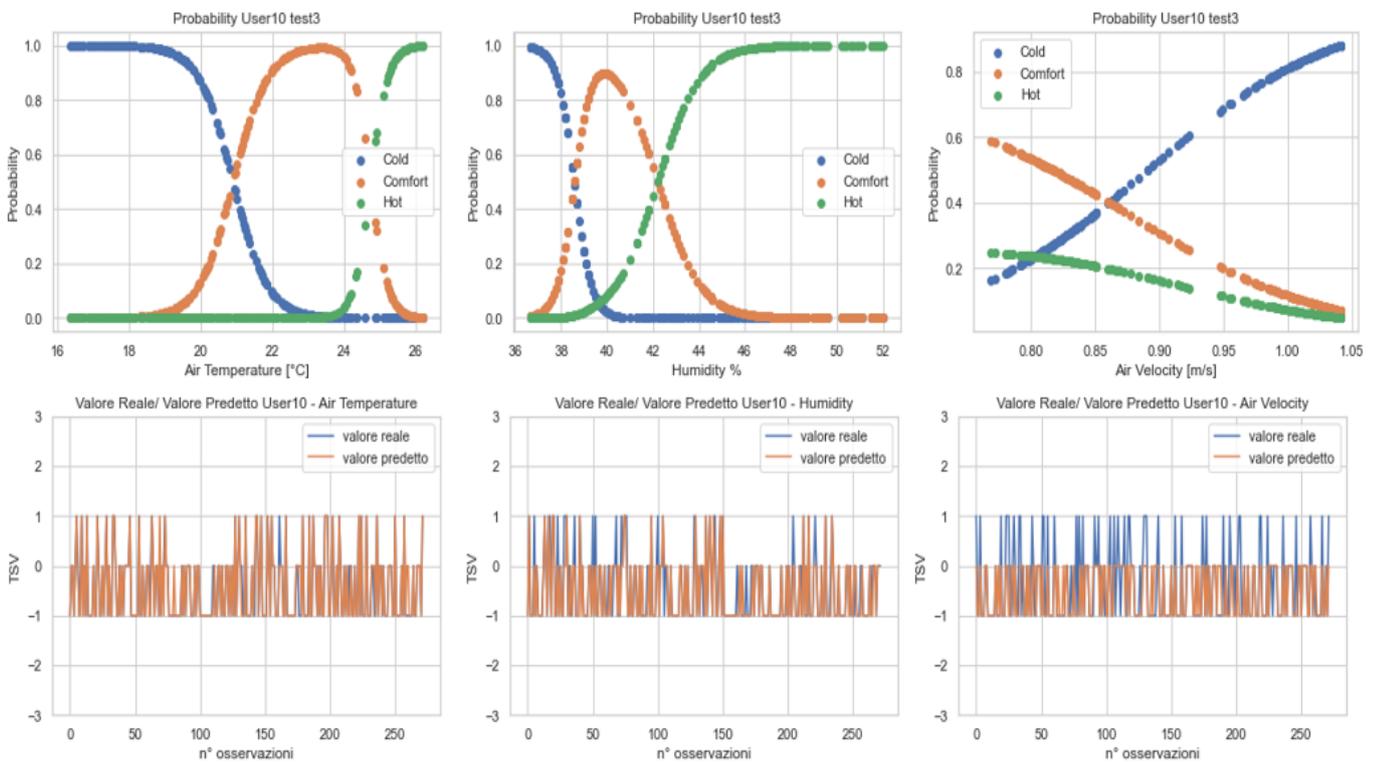


Figura 38 - Grafici di probabilità e di correlazione del soggetto 10

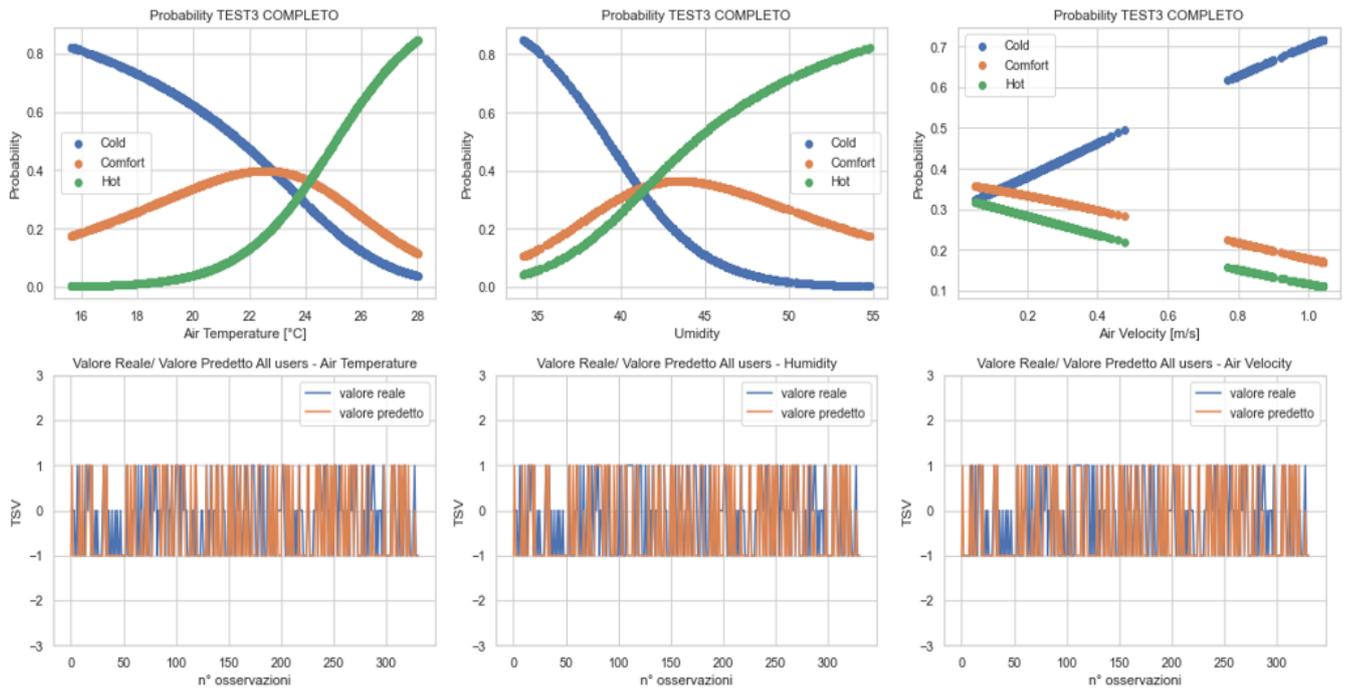


Figura 39 - Grafici di probabilità e di correlazione di tutti i soggetti

Si illustra ora come interpretare i grafici di probabilità. Facendo riferimento al grafico probabilistico del soggetto 4 (fig.34) con variabile di input t_a , ci si posiziona ad esempio a temperatura dell'aria pari a 20 °C. Per tale valore si osserva come la probabilità del soggetto 4 di essere in condizione di comfort è pari al 60 % (curva arancione), mentre quella relativa alla condizione di freddo è pari al 40 % e al caldo circa del 5/10 %. Ciò significa che l'algoritmo prevede che l'utente, per una temperatura dell'aria pari a 20 °C, abbia maggiore possibilità di trovarsi in condizione di benessere termico rispetto a quella di discomfort. Stessa procedura è stata utilizzata per tutti e 5 gli utenti e per tutte e 3 le variabili ambientali. È stato quindi possibile individuare delle zone con maggiore probabilità di comfort per ogni utente. Partendo dalla variabile t_a si possono osservare tali range di temperatura in tabella 2.

User	Range T_{air} [°C]	Probabilità Comfort %
4 9	20 - 22	60
6 10	22 - 24	80 - 100
8	24 - 26	60
All Users	22 - 24	40

Tabella 2 - Range di t_a con maggiore probabilità di comfort

Per quanto riguarda l'umidità dell'aria RH, si ha invece:

User	Range RH	Probabilità Comfort %
4	44 - 46	80
6	38 - 41	80
8	35 - 37	60/80
9	44 - 48	70
10	39 - 41	90
All Users	42 - 44	40

Tabella 3 - Range di RH con maggiore probabilità di comfort

Infine, per la velocità dell'aria, si ha:

User	Range Vair [m/s]	Probabilità Comfort %
4 9	0,10 – 0,15	30/35
6 8	0,05 – 0,10	45 - 65
10	0,70 – 0,80	55/60
Tutti gli user	0,10 – 0,20	30

Tabella 4 - Range di v_a con maggiore probabilità di comfort

Una volta analizzati i grafici di probabilità e ricavate le zone relative alla maggiore probabilità di comfort, è stata verificata la precisione dell'algoritmo rispetto a tutte e 3 le variabili di input, appunto attraverso il calcolo dell'indice di *correlazione di Pearson* e l'indice dell'*accuratezza*. Essi sono stati ricavati per tutti i soggetti relativamente a tutti e 3 i parametri ambientali. I valori sono stati raccolti all'interno della tabella 5.

USER	PEARSON			ACCURATEZZA		
	t_a	RH	v_a	t_a	RH	v_a
4	0.55	0.61	0.41	0,73	0,78	0,52
6	0.67	0.69	0.17	0,59	0,68	0,39
8	0.52	0.66	0.33	0,67	0,68	0,71
9	0.57	0.57	0.31	0,72	0,76	0,53
10	0.95	0.89	0.63	0,92	0,85	0,75
All Users	0.61	0.63	0.22	0,6	0,61	0,48

Tabella 5 - Coefficiente di Pearson e Accuratezza del test preliminare

4.2. Risultati del test con la termocamera

In questa seconda parte di studio si è utilizzato l'algoritmo Python per studiare l'andamento della percezione termica degli utenti al variare dei parametri fisiologici, raccolti a partire da immagini termografiche ottenute con la termocamera. Si è dunque allenata la Logistic Regression singolarmente sulle due variabili fisiologiche in esame, ovvero la temperatura della fronte T_{FRONTE} e la temperatura delle guance T_{GUANCE} . Anche in questo caso si è dunque lavorato su dataset distinti, uno relativo alla variabile di ingresso T_{FRONTE} e l'altro relativo a T_{GUANCE} . Come nel caso precedente si sono quindi ottenute due tipologie di grafici, una che mostra la probabilità di appartenenza ad una delle 3 classi (freddo, comfort, caldo) e l'altra che evidenzia l'andamento del TSV reale e predetto dall'algoritmo. Al termine si è poi verificata la precisione dello stesso tramite il calcolo dell'indice di *correlazione di Pearson* e l'*accuratezza*, anche in questo caso per ognuna delle due variabili fisiologiche di input.

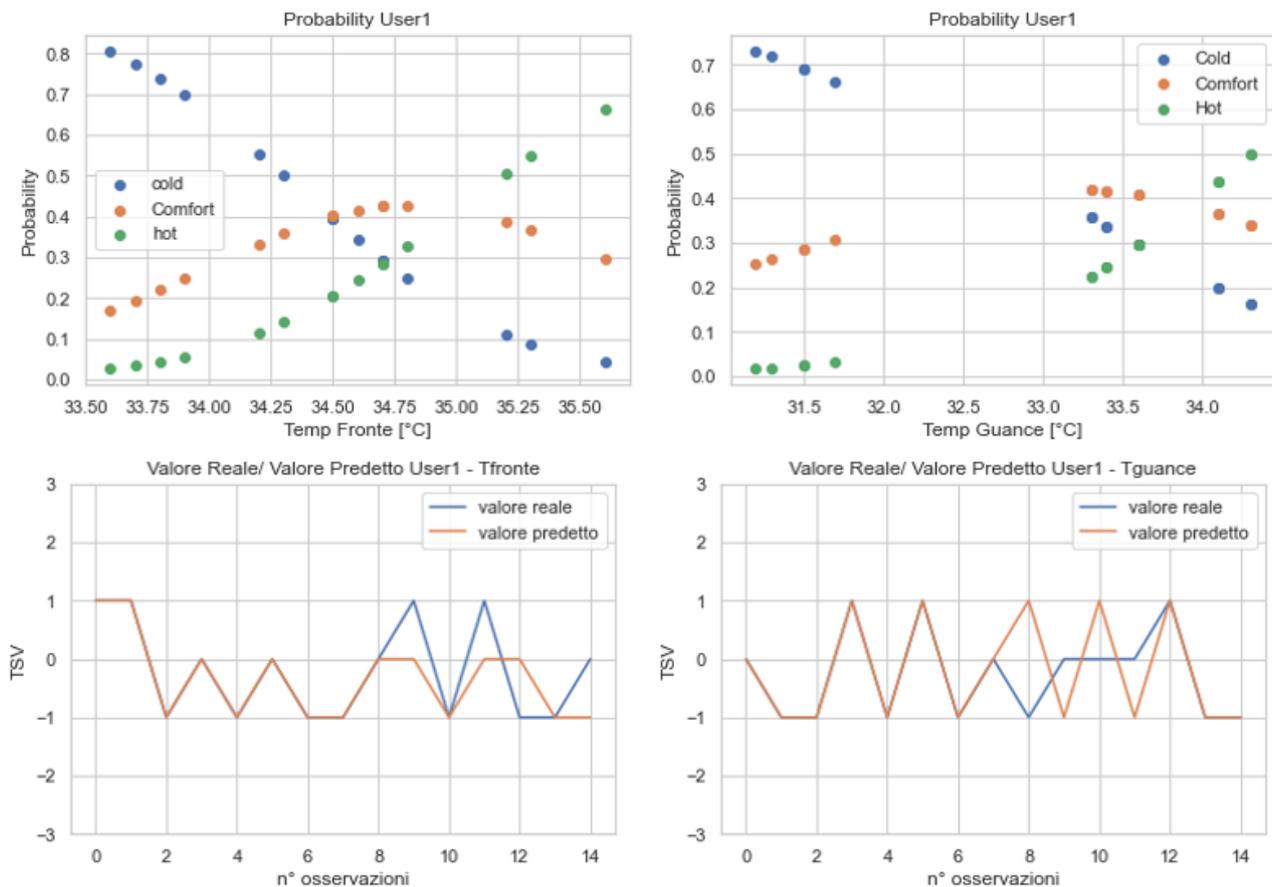


Figura 40 - Grafici di probabilità e di correlazione del soggetto 1

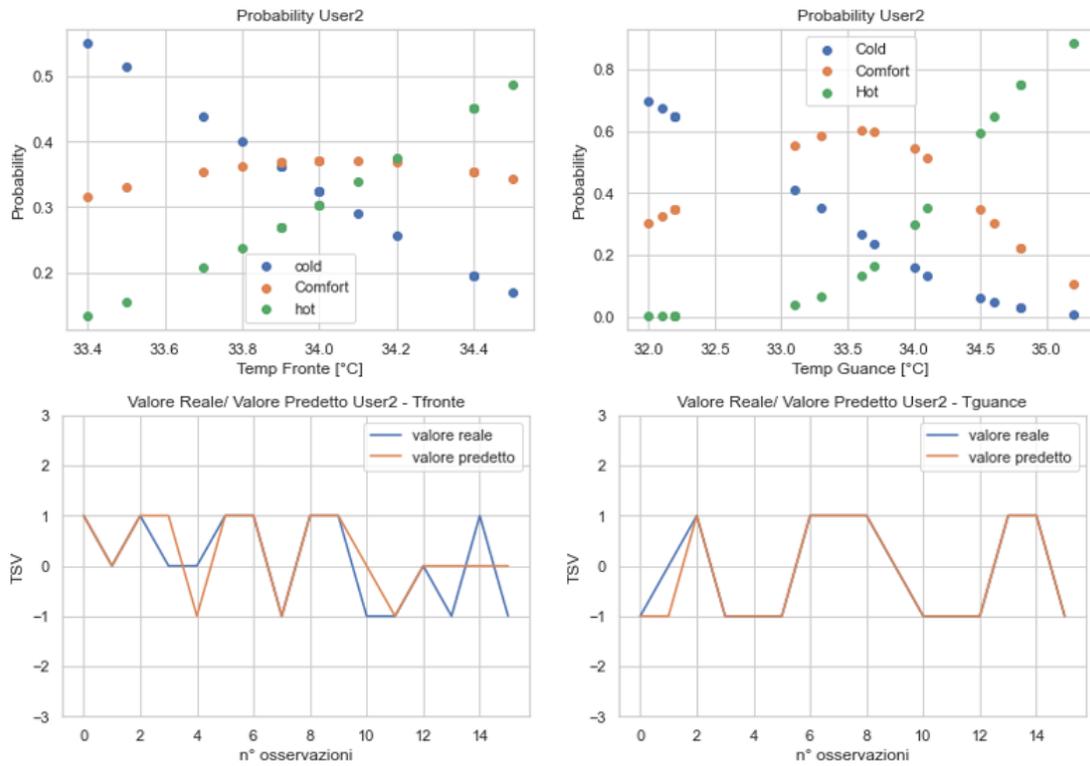


Figura 41 - Grafici di probabilità e di correlazione del soggetto 2

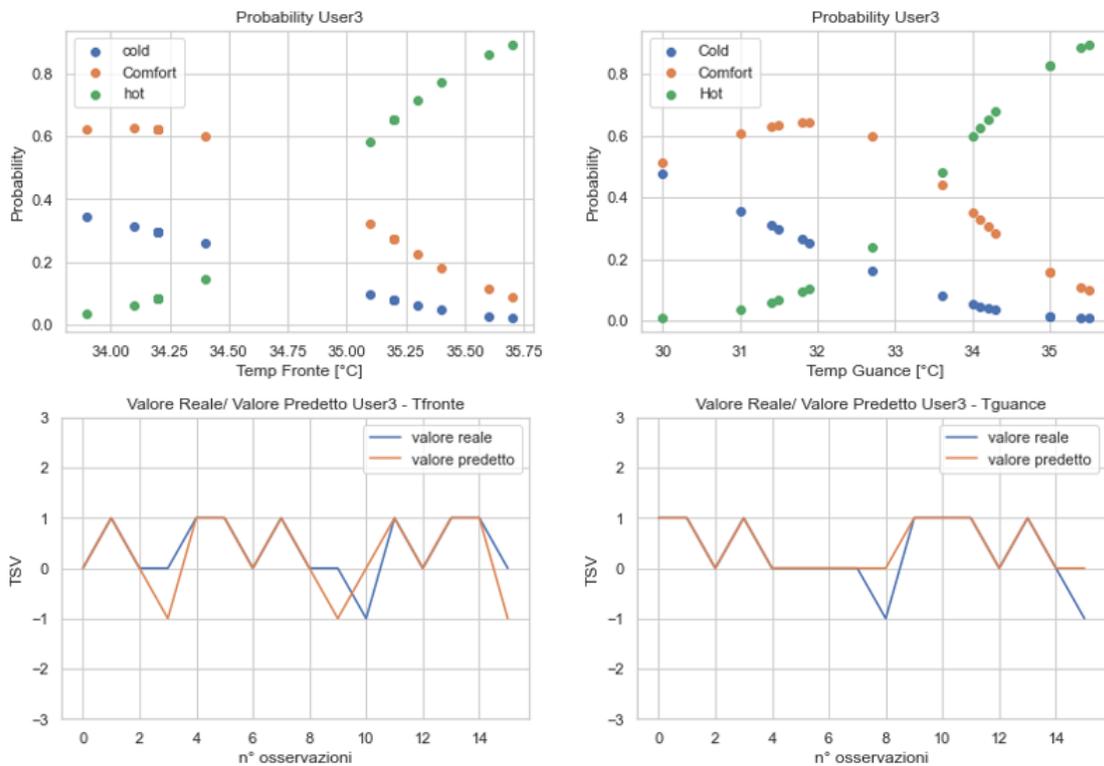


Figura 42 - Grafici di probabilità e di correlazione del soggetto 3

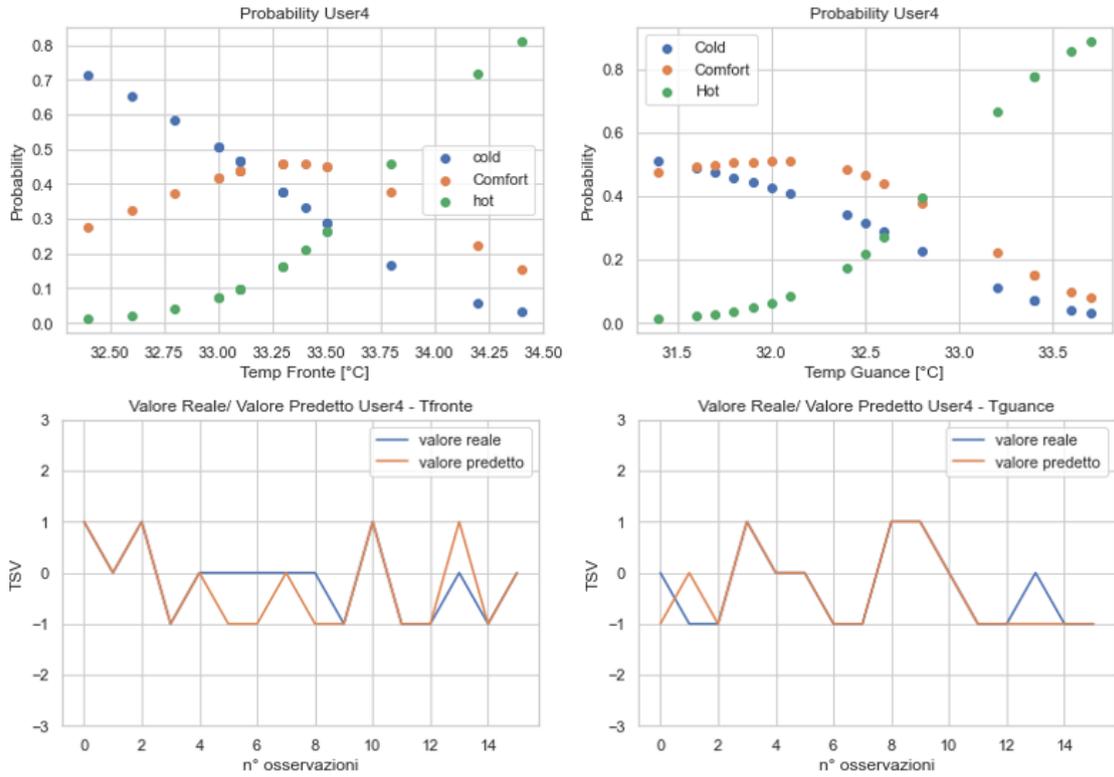


Figura 43 - Grafici di probabilità e di correlazione del soggetto 4

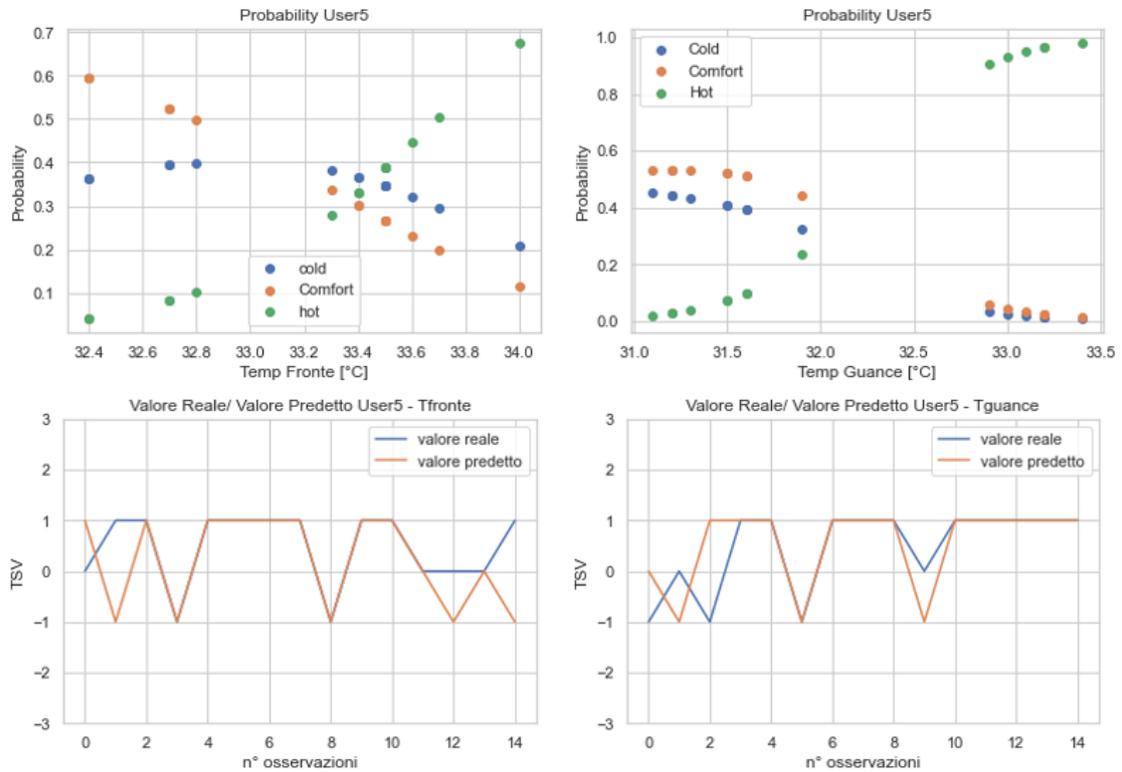


Figura 44 - Grafici di probabilità e di correlazione del soggetto 5

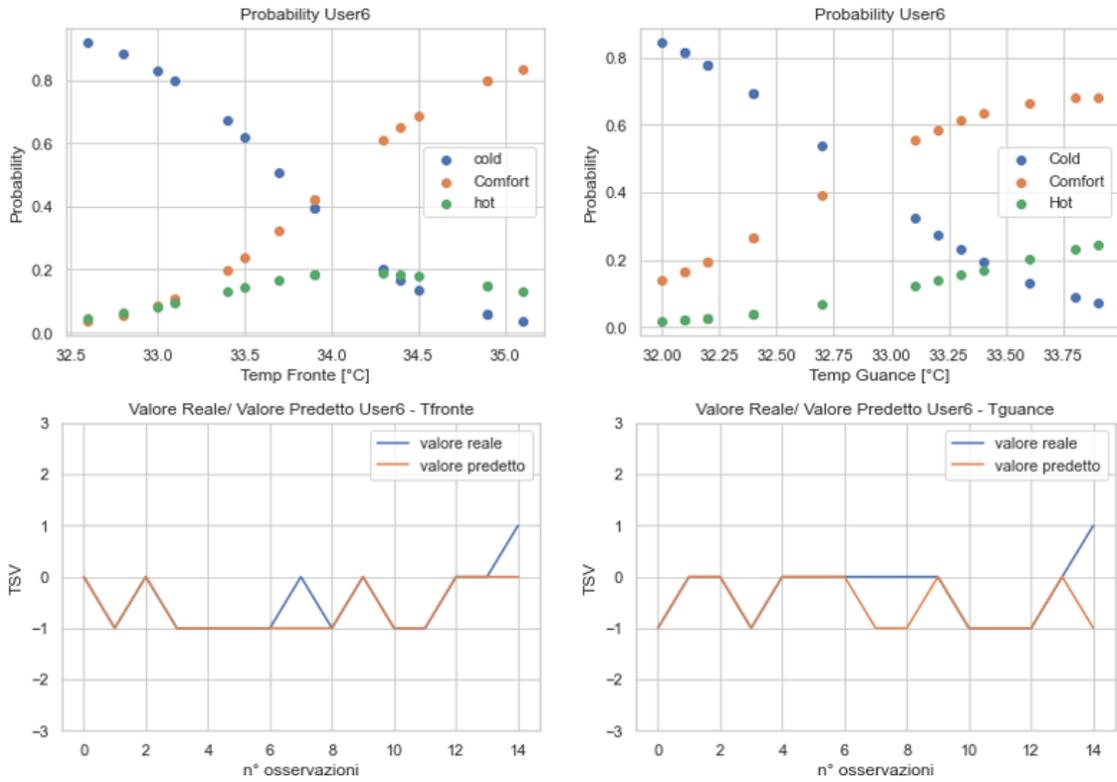


Figura 45 - Grafici di probabilità e di correlazione del soggetto 6

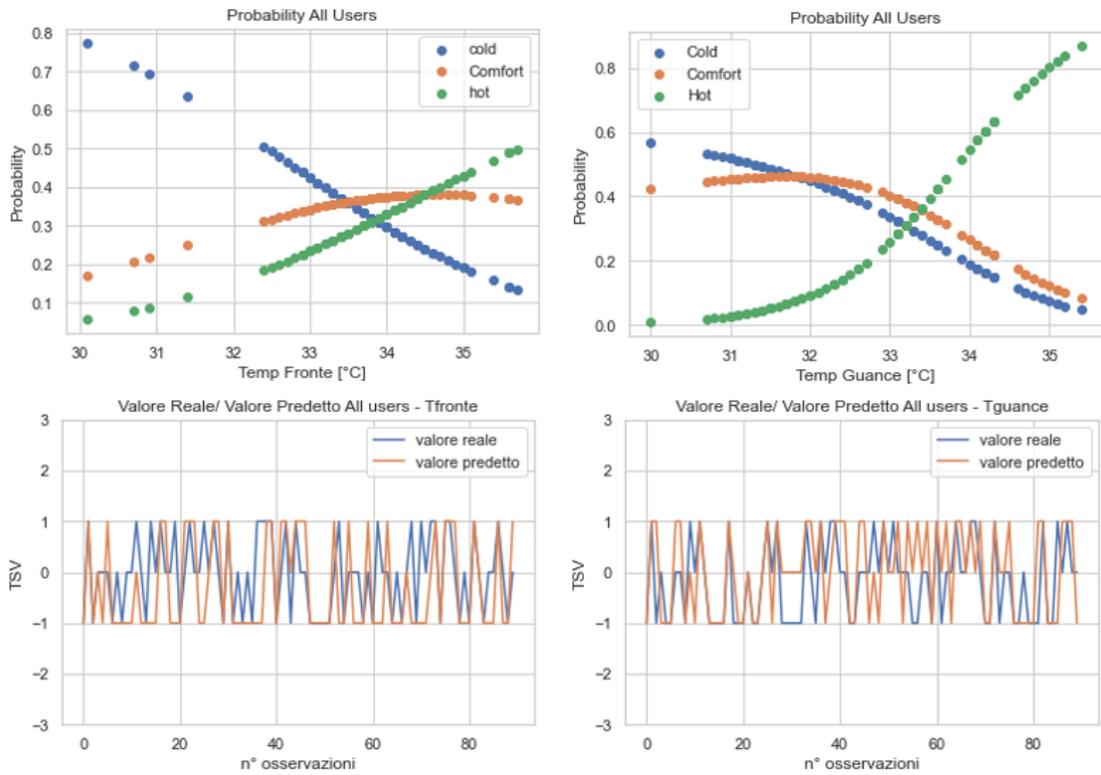


Figura 46 - Grafici di probabilità e di correlazione di tutti i soggetti

Anche in questo caso si sono potute individuare zone di T_{FRONTE} e di T_{GUANCE} nelle quali si riscontra una maggiore probabilità di comfort da parte dei soggetti. Questi range sono illustrati all'interno delle tabelle:

User	Range T_{FRONTE} [°C]	Probabilità Comfort %
1	34,6 – 35,2	40
2	33,9 – 34,2	40
3	34 – 34,5	60
4	33,2 – 33,5	50
5	32,4 – 32,6	60
6	34,5 – 35,1	80
All Users	34 - 35	40

Tabella 6 - Range di T_{FRONTE} con maggiore probabilità di comfort

User	Range T_{GUANCE} [°C]	Probabilità Comfort %
1	34,2 – 34,7	40
2	33,2 – 33,7	50
3	32,2 – 32,5	60
4	32,4 – 32,7	50
5	32,6 – 33	40
6	33,2 – 33,7	50
All Users	32,5 – 33,5	40

Tabella 7 - Range di T_{GUANCE} con maggiore probabilità di comfort

A questo punto è stata verificata la precisione dell'algoritmo rispetto alle 2 variabili di input, di nuovo attraverso il calcolo dell'indice di *correlazione di Pearson* e l'indice di *accuratezza* (tab.8).

USER	PEARSON		ACCURATEZZA	
	T_{FRONTE}	T_{GUANCE}	T_{FRONTE}	T_{GUANCE}
1	0,79	0,71	0,87	0,66
2	0,77	0,86	0,56	0,87
3	0,85	0,92	0,68	0,93
4	0,81	0,96	0,81	0,94
5	0,73	0,78	0,80	0,80
6	0,71	0,92	0,80	0,87
All Users	0,50	0,69	0,50	0,6

Tabella 8 - Coefficiente di Pearson e Accuratezza del test con la termocamera

5. CONCLUSIONI

Questo progetto di tesi aveva lo scopo di ottenere previsioni accurate del comfort termico a partire dai parametri ambientali e fisiologici misurati sull'utente. Tali previsioni possono mettere le basi (e serviranno) per sviluppare delle strategie di controllo capaci di garantire il benessere termico al più alto numero di utenti all'interno di un ambiente indoor.

Dall'analisi dei dati del test preliminare è stato possibile ottenere previsioni del comfort degli utenti al variare delle variabili ambientali quali la temperatura, l'umidità e velocità dell'aria, riuscendo a distinguere dei range di valori per ognuna delle 3 variabili, nei quali l'algoritmo individua la probabilità che un utente si trovi in comfort o in discomfort, a partire dai soli parametri ambientali. Per quanto riguarda la t_a si può osservare una zona comune a tutti gli utenti per valori compresi tra 22°C e 24°C dove si riscontra la probabilità più alta e pari al 40%. Per l'RH si ha invece un range comune a tutti gli utenti che oscilla tra il 42% e il 44% con una probabilità del 40%. Infine, per la v_a il range va da 0.10 m/s a 0.20 m/s con probabilità del 30%. Si noti però come tali valori corrispondano al caso in cui il modello viene allenato su tutti i soggetti, mentre, se analizzati singolarmente, le probabilità di comfort relative ai range individuati di t_a , RH e v_a sono sensibilmente superiori (tabelle 2,3,4).

Oltre all'individuazione delle fasce di comfort relative ai vari parametri ambientali, dall'osservazione dei valori di *accuratezza* e del coefficiente di *Pearson*, si può notare come la t_a e l'RH permettano di ottenere dei modelli predittivi migliori rispetto alla v_a , sia valutando gli utenti contemporaneamente, sia osservando i valori per i singoli soggetti.

Nella seconda parte dello studio invece, focalizzata sui test con la termocamera, è stato possibile ottenere previsioni del comfort degli utenti al variare dei parametri fisiologici quali la temperatura della fronte e delle guance. Si sono anche in questo caso individuate delle zone di comfort per entrambe le variabili di input, riscontrando che, analizzando contemporaneamente tutti gli utenti, il range di T_{FRONTE} con maggiore probabilità di benessere termico si attesta tra i 34°C e i 35°C con probabilità del 40%. Per la T_{GUANCE} , il range che massimizza la probabilità generale va dai 32.5°C a 33.5°C con una probabilità anche in questo caso del 40%. Analogamente alle considerazioni fatte per il test preliminare, se si osservano i singoli utenti le probabilità relative alle varie fasce di comfort individuate sono sensibilmente più elevate.

Per quanto riguarda l'efficacia dei modelli predittivi relativi ai due parametri fisiologici di input, osservando i valori del coefficiente di Pearson e dell'accuratezza si nota come la temperatura delle guance (T_{GUANCE}) fornisca un'accuratezza ed una correlazione maggiore sia nel caso si considerino contemporaneamente tutti i soggetti sia in quello in cui vengano analizzati singolarmente.

Lo scopo della tesi è stato dunque quello di indagare quali parametri ambientali e fisiologici permettano di ottenere una previsione migliore del comfort termico umano, con particolare attenzione ai parametri fisiologici. Dai test effettuati con la termocamera e dall'analisi degli indici di correlazione e di accuratezza degli algoritmi allenati tramite la T_{FRONTE} e la T_{GUANCE} , è stato possibile individuare nella T_{GUANCE} il parametro che meglio consente di prevedere il benessere termico umano (ovvero il TSV).

In un possibile progetto futuro potrebbe risultare interessante utilizzare l'algoritmo di controllo sviluppato nel lavoro di tesi per andare a "settare" un termostato programmabile. Esso potrebbe essere in grado di regolare il set-point dell'ambiente indoor grazie alle previsioni fornite dall'algoritmo al variare di parametri fisiologici (T_{GUANCE}), rilevabili da una termocamera.

6. BIBLIOGRAFIA

- [1] <https://www.studioessepi.it/magazine/ambiente/microclima-e-benessere-termico-normativa-e-indici-riferimento>
- [2] «A Better Way to Predict Comfort: The New ASHRAE Standard 55-2004», Agosto. <https://escholarship.org/uc/item/2m34683k>. Olesen, B. W., e G. S. Brager. 2004.
- [3] “Development and validation of a smart HVAC control system for multi-occupant offices by using occupants’ physiological signals from wristband” di Zhipeng Deng, Qingyan Chen, School of Mechanical Engineering, Purdue University, 585 Purdue Mall, West Lafayette, IN 47907, USA, Febbraio 2020
- [4] “Personalized human comfort in indoor building environments under diverse conditioning modes”, di Da Li, Carol C. Menassa*, Vineet R. Kamat, Dept. of Civil and Environmental Engineering, University of Michigan, United States, Ottobre 2017
- [5] “HEAT - Human Embodied Autonomous Thermostat”, Da Li^a, Carol C. Menassa^b, Vineet R. Kamat^b, Eunshin Byon^c, ^a Glenn Department of Civil Engineering, Clemson University, SC, 29634, United States, ^b Department of Civil and Environmental Engineering, University of Michigan, MI, 48109, United States, ^c Department of Industrial and Operations Engineering, University of Michigan, MI, 48109, United States, Maggio 2020
- [6] *Ciclo for*, https://it.wikipedia.org/wiki/Ciclo_for
- [7] <https://www.ai4business.it/intelligenza-artificiale/machine-learning/machine-learning-cosa-e-applicazioni/>
- [8] <https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471>
- [9] <https://lorenzogovoni.com/regressione-logistica/>
- [10] https://it.wikipedia.org/wiki/Indice_di_correlazione_di_Pearson
- [11] Spearman C., 1904. The proof and measurement of association between two things. American Journal of Psychology 15: 72–101.

[12] <https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9>

[13] <https://biblus.acca.it/focus/benessere-termoigrometrico-e-comfort-termico/#Modelli-di-comfort-termico>