



Ricerca di Sistema elettrico

Modelli di Machine Learning per disaggregazione dei consumi elettrici domestici e previsione del consumo elettrico

G. Oriolo, V. Piccialli, A. M. Sudoso

MODELLI DI MACHINE LEARNING PER LA DISAGGREGAZIONE DEI CONSUMI ELETTRICI DOMESTICI E PER LA PREVISIONE DEL CONSUMO ELETTRICO

G. Oriolo (Università degli Studi di Roma Tor Vergata), V. Piccialli (Sapienza Università di Roma), A. M. Sudoso (Università degli Studi di Roma Tor Vergata)

Gennaio 2022

Report Ricerca di Sistema Elettrico

Accordo di Programma Ministero dello Sviluppo Economico - ENEA

Piano Triennale di Realizzazione 2019-2021 – III annualità

Obiettivo: Tecnologie

Progetto: Tecnologie per la penetrazione efficiente del vettore elettrico negli usi finali

Work package: Local Energy District

Linea di attività: *(titolo completo)*

Responsabile del Progetto: Claudia Meloni, ENEA

Responsabile del Work package: Claudia Meloni, ENEA

Il presente documento descrive le attività di ricerca svolte all'interno dell'Accordo di collaborazione Disaggregazione dei consumi elettrici domestici tramite modelli di Machine Learning integrati con tecniche di previsione di consumo elettrico”

Responsabile scientifico ENEA: Claudia Snels

Responsabile scientifico DICII Università degli Studi di Roma Tor Vergata: Veronica Piccialli e da settembre 2021 Gianpaolo Oriolo.

Indice

1. SOMMARIO	4
INTRODUZIONE	5
DESCRIZIONE DELLE ATTIVITÀ SVOLTE E RISULTATI	6
1. ANALISI DEI DATI E PRE-PROCESSAMENTO	6
2. TECNICHE DI DEEP LEARNING PER LA DISAGGREGAZIONE DEI CONSUMI ELETTRICI.....	15
2.1 MULTI-TASK DEEP NEURAL NETWORK CON MECCANISMO DI ATTENTION	16
2.2 RISULTATI OTTENUTI TRAMITE LA RETE DEEP	19
3 CLUSTERING E PREVISIONE DEL CONSUMO COMPLESSIVO	22
3.1 LETTERATURA CORRELATA	23
3.2 GENERAZIONE DEI CLUSTER RELATIVI AI CONSUMI ELETTRICI	25
3.3 PREVISIONE DEI CONSUMI ELETTRICI	33
3.4 RISULTATI DELLA PREVISIONE	39
3.4.1 Previsioni a 5 minuti.....	40
3.4.2 Confronto tra previsioni LSTM ed Ensemble Learning.....	56
3.4.3 Previsioni a 15 minuti.....	59
3.4.4 Confronto tra previsioni LSTM ed Ensemble Learning.....	75
4 ANALISI DEI DATI DELLE NUOVE CASE FORNITI A NOVEMBRE 2021	77
4.1 RISULTATI DI BQP SUI NUOVI DATI	83
CONCLUSIONI	87
RIFERIMENTI BIBLIOGRAFICI.....	88

1. Sommario

Questo documento contiene la sintesi delle attività svolte all'interno dell'Accordo di collaborazione "Disaggregazione dei consumi elettrici domestici tramite modelli di Machine Learning integrati con tecniche di previsione di consumo elettrico". Sono stati affrontati con strumenti di machine learning avanzati il problema della disaggregazione dei consumi elettrici dei singoli elettrodomestici dato il consumo complessivo, il problema della costruzione di profili energetici, e infine quello della previsione dei consumi elettrici di un'abitazione con un dato profilo energetico. Tutti gli strumenti introdotti sono stati testati con successo su dati forniti da Enea relativi ad abitazioni residenziali a Roma e Viterbo. Combinando i diversi tool introdotti si arriva a una completa comprensione delle abitudini di consumo relativi alle abitazioni esaminate, che può essere utilizzata sia per migliorare le abitudini di consumo che per migliorare e personalizzare l'offerta di energia.

Introduzione

In questo lavoro ci si è focalizzati su tecniche per la previsione e il monitoraggio dei consumi elettrici in tutta la loro totalità, combinando due strumenti diversi: la disaggregazione dei consumi elettrici, che vuol dire individuare il contributo degli elettrodomestici energivori, e la previsione dei consumi totali di un'abitazione dopo averla inquadrata in una classe di consumi energetici. Tutta l'analisi è basata sull'utilizzo di tecniche avanzate di machine learning su dati storici di consumo elettrico di diverse abitazioni residenziali in cui vengono monitorati sia il consumo complessivo tramite smart meter che il consumo di alcuni elettrodomestici energivori tramite plug. Abbiamo integrato i dati storici estratti dal meter e dalle plug con dati ambientali (temperatura, pressione e umidità) e con dati comportamentali estratti da questionari compilati dai residenti nelle abitazioni monitorate. Le informazioni di diversa natura sono state combinate per ottenere previsioni robuste dei consumi elettrici complessivi di un'abitazione, e la disaggregazione del consumo tra gli elettrodomestici monitorati e quello che viene chiamato "sommerso", ovvero il consumo derivante da tutti gli appliance non monitorati. Per fare previsione dei consumi totali e per individuare correttamente le attivazioni e i consumi degli elettrodomestici abbiamo utilizzato tecniche di tipo supervisionato; data la previsione dei consumi e la sua disaggregazione, la differenza tra il totale previsto e la somma degli elettrodomestici monitorati rappresenta il sommerso, che può quindi essere dedotto nella sua totalità con una certa accuratezza. Sono state utilizzate due tecniche molto avanzate per la disaggregazione che sono in qualche modo ortogonali tra di loro: una rete deep proposta di recente specificatamente per il task di disaggregazione e un metodo basato sull'ottimizzazione di un problema di programmazione quadratica binario non convesso. L'approccio basato su reti deep è applicabile ogni qual volta si abbiano a disposizione buone quantità di dati storici. Il costo computazionale dell'addestramento è piuttosto elevato, e ha lo svantaggio di richiedere l'addestramento di una rete per ogni elettrodomestico di cui si vogliono predire i consumi, ma ottiene prestazioni molto elevate di accuratezza, e si può usare anche quando la percentuale di sommerso è elevata. Inoltre, è possibile addestrare un elettrodomestico su una casa e poi fare solo fine tuning su una casa con consumi simili. Se invece si hanno a disposizione pochi dati, si hanno molti elettrodomestici da disaggregare, e un sommerso non troppo elevato, il metodo basato sull'ottimizzazione è la strada da percorrere, in quanto sono sufficienti tre settimane di dati storici per stimare i parametri del modello di ottimizzazione, e, se si ha una granularità di dati all'ordine di minuti, il tempo di soluzione del problema di ottimizzazione tramite software standard è dell'ordine di pochi minuti. Inoltre, se il sommerso non è troppo elevato, si ottengono risultati ottimi in termini di accuratezza della previsione e di riconoscimento delle attivazioni dei diversi elettrodomestici. Nel corso del progetto ci sono stati forniti due dataset che ricadono esattamente nelle ipotesi di applicabilità dei due metodi e quindi abbiamo applicato entrambi con risultati molto buoni. Ci siamo poi concentrati sulla profilazione dei consumi e sulla previsione dei consumi complessivi. Un primo step è stato quello di classificare sulla base di informazioni combinate sulle abitudini di consumo e sui consumi medi degli elettrodomestici diversi profili energetici tramite tecniche non supervisionate di clustering. Poi, per ogni cluster, è stato costruito un modello di previsione dei consumi complessivi che può essere utilizzato su tutte le case appartenenti al cluster. Per costruire questo modello abbiamo combinato diverse tecniche di machine learning: gradient boosting, una rete neurale ricorrente e un classificatore atto a riconoscere i picchi di consumi. Combinando i due contributi, disaggregazione dei consumi e previsione del main si acquisisce una visione completa dei consumi di un'abitazione, e di abitazioni con comportamenti energetici simili.

Descrizione delle attività svolte e risultati

1. Analisi dei dati e pre-processamento

La fase di pre-processamento costituisce il primo passo fondamentale per una corretta analisi dei dati, in quanto la validità dei risultati dipende fortemente dalla qualità dei dati in input, in modo da evitare il noto problema del “garbage in, garbage out”.

Il dataset fornitoci inizialmente contiene i consumi elettrici di dieci case presenti nella periferia di Roma nel periodo compreso tra il 1 Gennaio 2019 e il 30 Giugno 2019. In una seconda fase del progetto sono stati aggiunti alcuni dati ma relativi solo a 3 case. I dati che sono stati forniti sono relativi ai diversi elettrodomestici, i cui consumi disaggregati sono stati rilevati tramite *smart plug*, e al sensore *energy meter* che ha misurato il consumo complessivo. L'*energy meter* misura sia la potenza istantanea che l'energia accumulata. Per quanto riguarda la potenza, il sensore invia un report ogni 5 secondi, o se viene rilevata una variazione del 10% della potenza o se questa è superiore a 50 Watt. Invece, per misurare l'energia il sensore invia un report nel momento in cui rileva una variazione in Watt del 10%.

In modo equivalente, anche gli smart plug hanno misurato sia la potenza istantanea che l'energia accumulata. Diversamente dall'*energy meter*, il sensore monitora la potenza inviando un report se il carico di alimentazione cambia dell'80%. Inoltre, le variazioni di potenza possono essere rilevate se il carico di corrente cambia del 15%. Per l'energia il sensore invia un report se vi è una variazione minima nel consumo di energia pari a 0.1 kWh.

Nella Tabella 1 sono riportati i dispositivi presenti in ogni casa. Il simbolo ✓ indica la presenza del dispositivo nella rispettiva abitazione.

	Casa									
	1	2	3	4	5	6	7	8	9	10
Frigo	✓	✓	✓	X	X	✓	✓	X	✓	✓
Lavastoviglie	✓	X	X	X	✓	✓	✓	X	✓	✓
Lavatrice	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Tv	✓	X	X	✓	X	X	X	✓	X	✓
Bimby	X	✓	X	X	X	X	X	X	X	X
Lampada	X	✓	X	X	X	X	X	X	X	X
Termoventilatore	X	✓	X	X	X	X	X	X	✓	X
Microonde	X	X	✓	X	X	✓	✓	X	✓	X
Aspirapolvere	X	X	X	✓	X	X	X	✓	X	X
Cucina	X	X	X	✓	X	X	X	X	X	X
Ferro da stiro	X	X	X	X	✓	X	X	X	X	X
Macchina caffè	X	X	X	X	✓	X	X	X	X	X

Tabella 1: Elettrodomestici presenti nelle diverse abitazioni

Poiché si è scelto di avere i dati ogni secondo, le osservazioni al millisecondo sono state arrotondate al secondo più vicino e nel caso in cui in un secondo ci fossero osservazioni duplicate, la potenza è stata sostituita con la media delle potenze delle due osservazioni. Tali operazioni sono state eseguite

sul dataset di ogni casa, in modo da poter procedere nelle fasi successive solo con i dati di interesse. In seguito, è stato quindi opportuno analizzare il dataset così ricostruito in modo più dettagliato, così da comprendere quanti dati si avessero a disposizione per ogni casa. In tal modo sono state realizzate delle heatmap.

In generale, una mappa di calore (o heatmap) è una rappresentazione visiva dei dati, dove i valori contenuti all'interno di una matrice sono rappresentati da colori. Nel caso del dataset in esame, tali heatmap sono state realizzate sia per i dati relativi al meter che per i dati di ogni elettrodomestico.

Si è quindi proceduto alla rappresentazione delle mappe mettendo lungo l'asse orizzontale il giorno e lungo l'asse verticale l'ora, facendo ciò per ogni mese presente. In tal modo, si è potuto vedere quali fossero i periodi con dati mancanti e come fossero distribuiti i valori delle potenze. Riportiamo come esempio nelle figure 1 e le heatmap ottenute per il frigo e per il meter di casa 10. Per gli altri dispositivi di tale casa sono state ottenute delle mappe simili a quelle del frigo. In particolare, i valori riportati con il colore verde acqua corrispondono ai periodi mancanti.

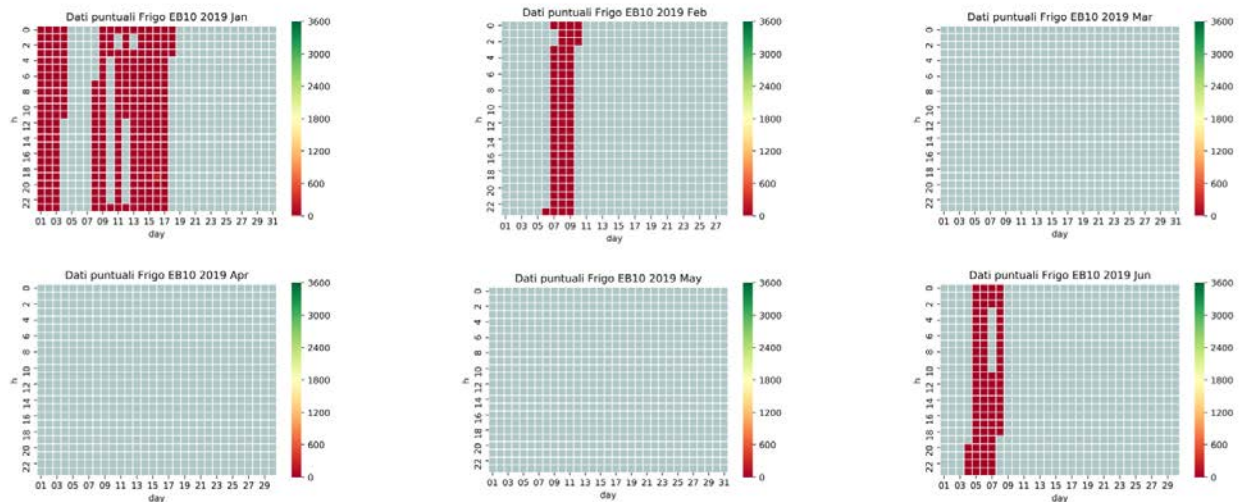


Figura 1: Heatmap del frigo di casa 10

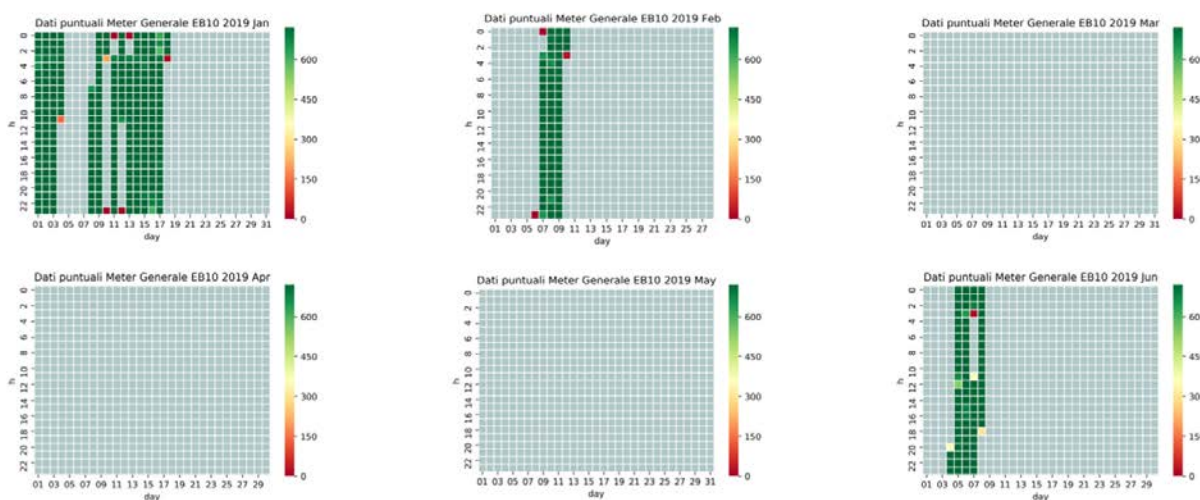


Figura 2: Heatmap dell'energy meter di casa 10

Una prima analisi delle heatmap ci ha portato a scartare dalle fasi successive alcune case, data la disponibilità di una quantità molto limitata di dati. In particolare, sono stati scartati i dati relativi al meter e agli elettrodomestici di casa 10, in quanto, come si può notare dalle Figure 1 e 2, nel dataset sono presenti rilevazioni relative solamente a qualche giorno di gennaio, febbraio e giugno.

Inoltre, sono stati eliminati anche tutti i dati dei dispositivi di casa 2, in quanto risultava disponibile solo il mese di gennaio; tuttavia, per tale casa, è stato comunque salvato il dataset relativo al meter perché utile per la successiva fase di previsione del consumo complessivo. Infine, sono stati scartati anche i seguenti dataset: tv di casa 8, aspirapolvere di casa 4, ferro da stiro di casa 5 e lavatrice di casa 3, poiché nelle heatmap realizzate si è notata una presenza molto frammentaria di dati, come si può anche notare dalla figura 3 nella quale sono riportate le heatmap ottenute per la lavatrice di casa 3.

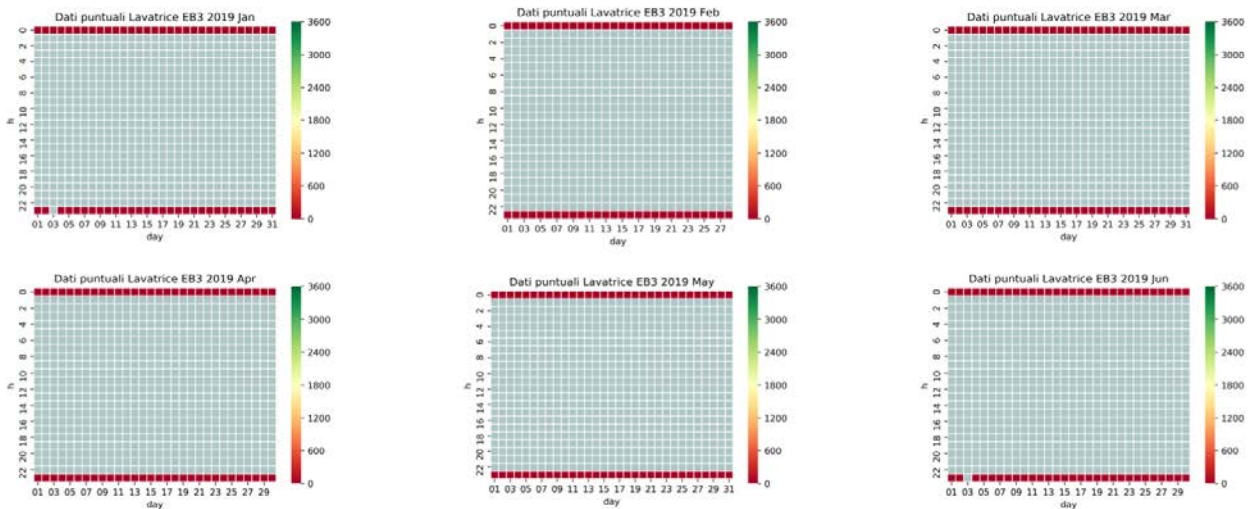


Figura 3: Heatmap Lavatrice casa 3

E' stato necessario vedere quali e quanti dispositivi si avessero a disposizione, in modo da decidere quali analizzare nelle fasi successive. Per fare ciò si è calcolata la frequenza totale di ogni elettrodomestico considerato. Le frequenze sono riassunte nella Tabella 2. In tal modo, si è deciso di scartare tutti gli elettrodomestici che avessero una frequenza minore o uguale a uno. Di conseguenza, i dispositivi salvati in tale fase sono: frigo, lavastoviglie, lavatrice, tv e microonde. In particolare, nella Tabella 3, sono riportati gli elettrodomestici che si hanno a disposizione per ogni casa.

	EB1	EB3	EB4	EB5	EB6	EB7	EB8	EB9	Frequenza
Frigo	✓	✓	✗	✗	✓	✓	✗	✓	5
Lavastoviglie	✓	✗	✗	✗	✓	✓	✓	✓	5
Lavatrice	✓	✗	✓	✓	✓	✓	✓	✓	7
Tv	✓	✗	✓	✗	✗	✗	✗	✗	2
Bimby	✗	✗	✗	✗	✗	✗	✗	✗	0
Lampada	✗	✗	✗	✗	✗	✗	✗	✗	0
Termoventilatore	✗	✗	✗	✗	✗	✗	✗	✓	1
Microonde	✗	✓	✗	✗	✓	✓	✗	✓	4
Aspirapolvere	✗	✗	✗	✗	✗	✗	✓	✗	1
Cucina	✗	✗	✓	✗	✗	✗	✗	✗	1
Ferro da stiro	✗	✗	✗	✗	✗	✗	✗	✗	0
Macchina caffè	✗	✗	✗	✓	✗	✗	✗	✗	1

Tabella 2 : Frequenza dei diversi elettrodomestici

	Frigo	Lavastoviglie	Lavatrice	Microonde	Tv
Casa 1	✓	✓	✓	✗	✓
Casa 3	✓	✗	✗	✓	✗
Casa 4	✗	✗	✓	✗	✓
Casa 5	✗	✓	✓	✗	✗
Casa 6	✓	✓	✓	✓	✗
Casa 7	✓	✓	✓	✓	✗
Casa 8	✗	✗	✓	✗	✗
Casa 9	✓	✓	✓	✓	✗

Tabella 3 : Elettrodomestici considerati in ogni casa

Le heatmap descritte sopra sono state utili anche per verificare la coerenza dei dati mancanti tra il meter e i vari dispositivi presenti nelle case in esame, in quanto ci si aspetterebbe che in assenza di dati del singolo dispositivo, non sia avvenuta una registrazione della potenza neanche da parte del meter. Quindi, analizzando le heatmap, sono stati individuati i periodi di incoerenza tra i dati del meter e dei dispositivi presenti, considerando, per ogni casa, gli elettrodomestici riportati nella tabella 3. Nella Tabella 4 è riportato con ✓ il caso in cui si verifica coerenza tra il meter della relativa casa e i dispositivi presenti, mentre con x il caso contrario.

Ad esempio, come si può vedere dalle heatmap relative agli elettrodomestici e al meter di casa 4, riportate nelle Figure 4, 5 e 6, in tutti i mesi presenti non vi è una corrispondenza dei dati mancanti tra il meter e i dispositivi. Un ragionamento analogo è stato effettuato anche per le altre case, ottenendo i risultati riportati nella Tabella 4.

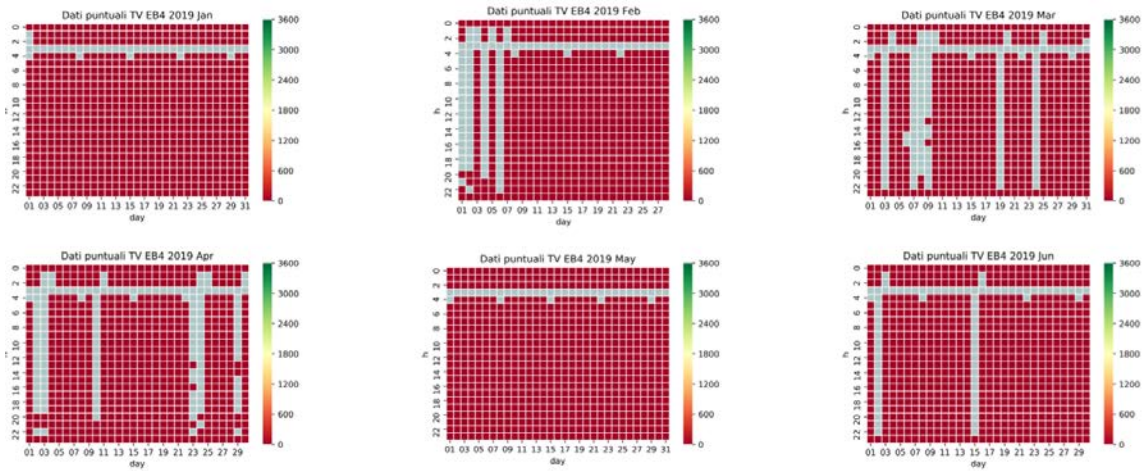


Figura 4: Heatmap della TV di casa 4

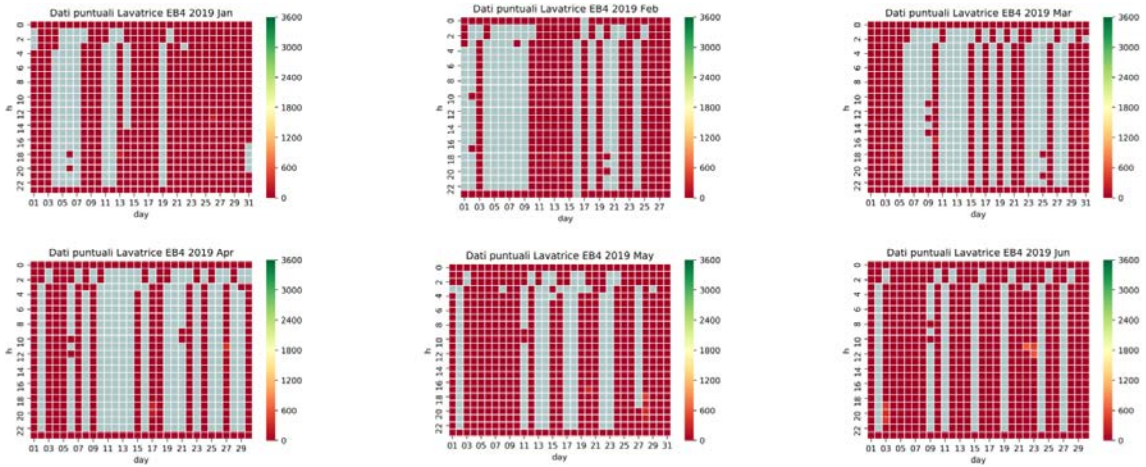


Figura 5: Heatmap della lavatrice di casa 4

Heatmap Lavatrice casa 3

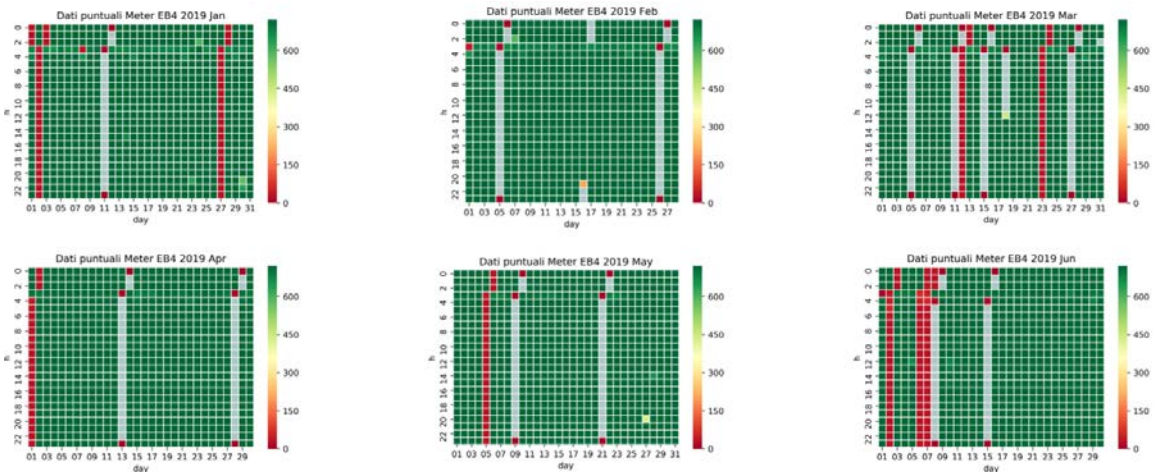


Figura 6: Heatmap del meter di casa 4

	EB1	EB3	EB4	EB5	EB6	EB7	EB8	EB9
Gennaio	✓	✗	✗	✓	✓	✓	✗	✓
Febbraio	✗	✗	✗	✓	✓	✓	✗	✓
Marzo	✓	✗	✗	✓	✓	✓	✓	✓
Aprile	✗	✗	✗	✓	✗	✓	✓	✗
Maggio	✓	✗	✗	✓	✗	✗	✓	✗
Giugno	✗	✗	✗	✓	✓	✗	✗	✗

Tabella 4: Coerenza tra meter e dispositivi

Tale analisi ha permesso di scartare tutti i dispositivi di casa 3 e casa 4, salvando solamente i relativi meter, in quanto, come si può osservare dalla tabella 4, in tutti i mesi vi è un'assenza di coerenza tra i dati degli elettrodomestici e del meter. Inoltre, avendo scartato i dispositivi di casa 4, come si può osservare nelle Tabelle 2 e 3, dopo le varie analisi la tv è presente solo nel dataset di casa 1, in quanto ritenuto inutilizzabile per la casa 4. Di conseguenza, si è deciso di scartare la tv anche per la casa 1, in quanto presente solo in tale abitazione.

	Frigo	Lavastoviglie	Lavatrice	Microonde
Casa 1	✓	✓	✓	✗
Casa 5	✗	✓	✓	✗
Casa 6	✓	✓	✓	✓
Casa 7	✓	✓	✓	✓
Casa 8	✗	✗	✓	✗
Casa 9	✓	✓	✓	✓

Tabella 5: Dispositivi considerati per ogni casa

In conclusione, nella Tabella 5 sono riportati i dispositivi che sono stati salvati per ogni casa. Per quanto riguarda casa 2, casa 3 e casa 4, si è deciso di salvare per le fasi seguenti solo i dati relativi al meter, in quanto utili per la fase di previsione. Per le diverse analisi abbiamo avuto bisogno di riportare il dataset a due diverse granularità: alla granularità di un secondo e alla granularità di un minuto, a seconda del task di machine learning da eseguire. In particolare, per effettuare la disaggregazione dei consumi tramite rete deep servono dati alla granularità di secondi, mentre per la previsione sono sufficienti dati aggregati al secondo.

L'obiettivo della fase di pre-processing è ottenere un dataset al secondo, avendo inizialmente a disposizione osservazioni ogni 5 secondi rilevate dal meter e osservazioni ad attivazione (soglia) rilevate dai vari elettrodomestici. Dopo aver effettuato le operazioni iniziali sul dataset, si è potuto procedere alla ricostruzione dei dati considerando, per ogni casa, solamente il meter e gli elettrodomestici riportati nella Tabella 5.

Come si è potuto notare dalle heatmap, le potenze in molti istanti temporali non sono state rilevate, quindi per evitare di avere periodi senza osservazioni, si è deciso di propagare l'ultima potenza nota, introducendo un lieve errore nel caso in cui l'intervallo mancante fosse piccolo. In tal modo si è ottenuto un dataset al secondo senza dati mancanti. Nella fase di preprocessing è stato utile realizzare dei plot per visualizzare l'andamento dei dati. Di conseguenza, dopo aver ricostruito il dataset al secondo, sono stati generati i plot per ogni casa, sia per i vari dispositivi che per il meter, ad intervallo temporale di un mese. Tali plot hanno permesso di identificare il pattern di ogni elettrodomestico e del meter. Come esempio nella Figura 7 sono riportati i plot ottenuti per i vari elettrodomestici e per il meter di casa 6, prendendo in considerazione il mese di febbraio dispositivi

come la lavastoviglie, la lavatrice e il microonde hanno dei pattern ad attivazione, mentre nel frigo è presente un pattern ad onda quadra, come si può anche evincere dall'ultimo plot sulla destra, nel quale è riportato l'andamento giornaliero di tale elettrodomestico.

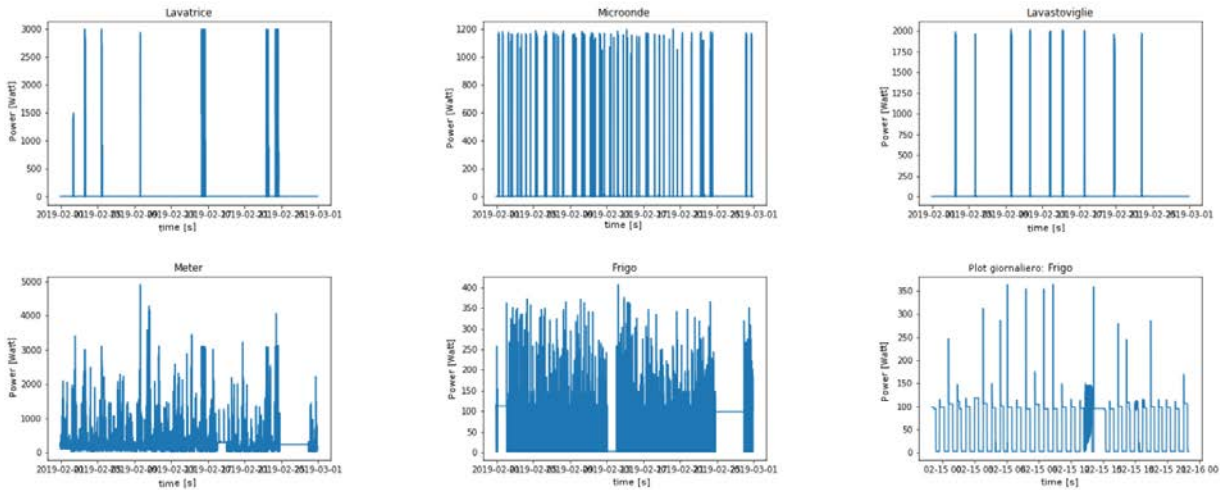


Figura 7: Plot dei diversi dispositivi e del meter di casa 6 nel mese di febbraio

Dall'analisi dei plot dei dati a disposizione è stata rilevata la presenza di outliers in due casi, ovvero nel frigorifero della casa 9 e nella lavatrice della casa 5. Tali anomalie sono state sostituite utilizzando il metodo dell'interpolazione lineare descritto dalla seguente formula:

$$f(x) = \frac{f(x - 1) + f(x + 1)}{2}$$

Quindi il valore anomalo dell'istante x stato sostituito con la media tra il valore precedente e il successivo. Infine, è stato effettuato un controllo di coerenza tra meter ed elettrodomestici le potenze rilevate dal meter dovrebbero essere almeno pari alla somma delle potenze raggiunte dai dispositivi. Si è quindi verificato che la potenza del meter non fosse minore della somma delle potenze dei singoli dispositivi. Per ogni casa e per ogni mese sono state quindi calcolate le percentuali sul numero di volte in cui ci fosse inconsistenza, come il rapporto tra numero di inconsistenze (numero di istanti temporali in cui la potenza del meter fosse inferiore alla somma delle potenze dei singoli dispositivi) e totale di osservazioni.

	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Totale
Casa 1	0.012840	0.014300	0.013778	0.014875	0.008635	-	0.012907
Casa 5	0.026784	0.010607	0.0	0.013028	0.021336	0.077015	0.024807
Casa 6	0.020134	0.026878	0.022431	0.033203	0.051932	-	0.027698
Casa 7	0.024186	0.022284	0.021739	0.025206	0.025794	-	0.023718

Casa 8	0.004898	0.004001	0.004817	0.000709	0.000813	0.003674	0.003149
Casa 9	0.011629	0.014922	0.012528	0.009394	-	-	0.012914

Tabella 6: Percentuali di inconsistenza

Nella Tabella 6 sono riportate le percentuali che sono state ottenute per ogni casa e per ogni mese. Nella colonna totale le percentuali sono invece calcolate su tutto il periodo di riferimento; quindi i valori in tale colonna rappresentano il numero di volte in cui si verifica inconsistenza sul numero di osservazioni in tutto il periodo a disposizione. Infine, i casi indicati con il "-" si riferiscono ai mesi non presenti nel dataset della rispettiva casa. Come si può notare, i valori ottenuti sono molto bassi, quindi i casi in cui si ha inconsistenza tra meter e device sono rari. Per questo motivo, si è deciso di rimediare a tali errori sostituendo la potenza rilevata dal meter con la somma delle potenze dei dispositivi.

In generale, sono stati individuati i possibili casi di inconsistenza, trattando ognuno di questi in modo differente. Un primissimo caso è il caso in cui la somma delle potenze dei dispositivi è un valore maggiore del valore della potenza rilevata dal meter: in tale situazione si è sostituita la potenza del meter con la somma delle potenze rilevate dai singoli elettrodomestici. In altri casi, invece, il meter potrebbe non aver rilevato una potenza avendo però un valore concreto per la somma delle potenze dei dispositivi o viceversa. In queste situazioni la potenza del meter è stata lasciata con il valore effettivo o il NaN già presente. Da un'analisi visiva dei plot si è notato come, nel caso di intervalli di lunga durata, la propagazione al secondo dell'ultima potenza nota abbia generato dei tratti costanti, i quali avrebbero potuto causare problemi nelle successive fasi di disaggregazione e di previsione. Inoltre, si è controllato che i tratti costanti fossero coerenti tra il meter e i vari dispositivi presenti nella casa in esame. Per superare tale problema, analizzando i plot e le heatmap, sono stati estrapolati tutti i periodi con tale problema e nel dataset sono stati sostituiti con il valore NaN. Un esempio di tale problematica si può notare nei plot riportati nella figura 4, in particolare nel frigo e nel meter. Quindi, nei dati nei quali si è notato un problema analogo si è proceduto alla sostituzione con i valori NaN, che sono poi sostituiti con opportune medie dei periodi precedenti e successivi. In particolare, nel caso in cui ci fosse un tratto lungo x secondi con valori NaN, il valore di ogni secondo viene sostituito con la media tra il rispettivo valore presente nei precedenti x secondi e il valore dei successivi x secondi. Inoltre, se non fosse possibile calcolare la media, ovvero se il valore dei successivi x secondi fosse NaN o il periodo precedente non fosse presente, per evitare di introdurre ulteriore rumore nel dataset, si è scelto di sostituire il NaN con il valore dell'ultima potenza nota prima del tratto costante. In tal modo il valore della potenza è quello che vi era prima di mettere tutto il tratto a NaN, in quanto nello step precedente si era propagata l'ultima potenza nota per la ricostruzione del dataset al secondo.

Infine, si è visto che nel dataset di ogni casa molto spesso non vi era corrispondenza nel periodo di fine tra gli elettrodomestici e il rispettivo meter; per risolvere tale problema, per ogni casa sono stati allineati i periodi di inizio e fine tra tutti i dati a disposizione. Nella Tabella 7 sono riportate le date di inizio e fine per ogni casa.

	Data inizio	Data fine
CASA 1	2019-01-01 00:00:00	2019-05-29 18:59:00
CASA 2	2019-01-01 00:00:00	2019-02-02 19:29:00
CASA 3	2019-01-01 00:00:00	2019-05-27 19:10:00
CASA 4	2019-01-01 00:00:00	2019-06-30 23:59:00
CASA 5	2019-01-01 00:00:00	2019-06-30 23:59:00
CASA 6	2019-01-01 00:00:00	2019-05-11 13:03:00
CASA 7	2019-01-01 00:00:00	2019-05-21 09:21:00
CASA 8	2019-01-01 00:00:00	2019-06-30 23:59:00
CASA 9	2019-01-01 00:00:00	2019-04-02 06:04:00

Tabella 7: Date di inizio e fine per ogni casa

Dato il dataset a granularità al secondo, si è proceduto per aggregare i dati al minuto, dati utilizzati poi per la fase di previsione del consumo totale (main). Il dataset ricostruito al secondo è stato aggregato al minuto facendo la media dei dati presenti nel relativo minuto. Nella Figura 8 sono riportati i plot dei dati relativi al consumo complessivo della casa 1 e della casa 6, per i dati aggregati al minuto.

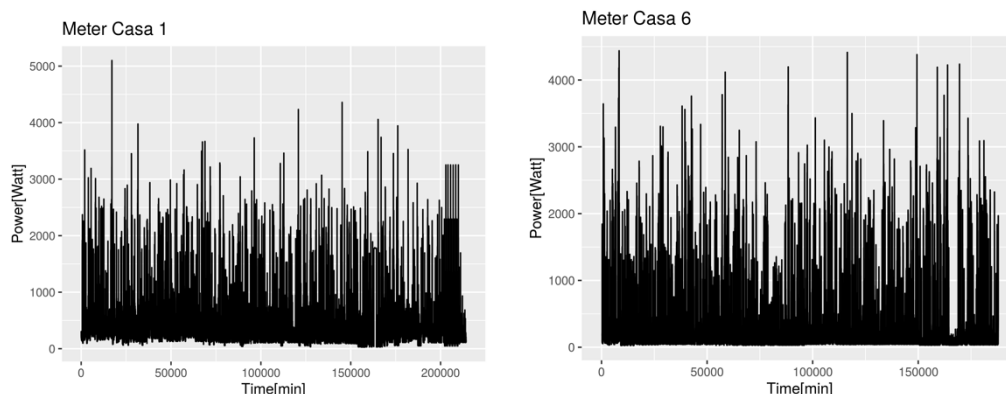


Figura 8: Meter aggregato al minuto per casa 1 e casa 6

I due dataset al secondo e al minuto ottenuti sono stati poi scalati. Sono state testate diverse tecniche di scaling e la più efficace è risultata quella Min-Max Scaler con range tra 0 e 1, poiché si sono ottenuti risultati migliori rispetto ad altri range di scaling e alla standardizzazione.

A settembre 2020 ci sono stati poi forniti dei dati aggiuntivi per le case C3, C8 e C9: per la casa C3 sono stati aggiunti dati dal 01/06/2019 al 01/04/2020, per la casa C8 dati dal 01/06/2019 al 15/01/2020, per la casa C9 dal 01/06/2019 al 06/12/2019. Questi dati sono stati preprocessati in modo analogo a quanto fatto per il dataset precedente, e sono stati inclusi negli esperimenti.

2. Tecniche di Deep Learning per la disaggregazione dei consumi elettrici

La crescente preoccupazione di migliorare l'efficienza energetica ha portato alla creazione di nuove tecnologie e obiettivi, come quelli relativi alla gestione dell'energia per controllare i picchi di domanda energetica e lo sviluppo di elettrodomestici sempre più efficienti. Tutti questi obiettivi possono essere raggiunti attraverso la progettazione e l'implementazione di algoritmi di disaggregazione dei consumi elettrici. La disaggregazione del carico è il processo di separazione del carico elettrico totale di una singola famiglia in carichi specifici dell'apparecchio. Questo può essere fatto sia attraverso il monitoraggio intrusivo del carico (ILM) o utilizzando tecniche di monitoraggio non intrusivo del carico (NILM). In un approccio intrusivo, è necessario collegare i contatori di energia ad ogni presa di corrente, il che porta ad alti costi di installazione e problemi di privacy. D'altra parte, utilizzando tecniche non intrusive, il consumo di energia dei principali elettrodomestici è stimato utilizzando solo la domanda di energia aggregata registrata da un singolo contatore elettrico che monitora più apparecchi. Il NILM rappresenta un'alternativa conveniente alle tecniche di monitoraggio intrusivo ed è riconosciuta come una tecnica chiave per abilitare servizi innovativi di smart grid. Al giorno d'oggi, gli smart meter dotati di tecnologia NILM sono ampiamente utilizzati dalle aziende per monitorare gli usi dell'energia elettrica in diverse abitazioni. Oltre a misurare i dati di consumo elettrico, i contatori intelligenti possono comunicare i dati raccolti ai fornitori di servizi e quindi consentire la creazione di nuovi servizi energetici come l'ambient-assisted living, la previsione della domanda, o raccomandazioni su come risparmiare energia. Una sfida fondamentale per la ricerca è come esplorare e sfruttare appieno il contenuto informativo dei dati dei contatori intelligenti.

Come per i problemi di apprendimento automatico correlati, la ricerca e lo sviluppo richiedono una quantità sufficiente di dati per addestrare e convalidare nuovi algoritmi. Guardando gli sforzi più recenti su NILM, le reti neurali deep hanno raggiunto prestazioni impressionanti in termini di precisione, ma la mancanza di dati supervisionati che consistono in dati energetici aggregati e sub-misurati per ogni casa monitorata è uno dei problemi principali che bloccano il miglioramento degli algoritmi di disaggregazione.

Dal punto di vista dell'apprendimento automatico, il NILM può essere visto come un problema di apprendimento multi-task composto da un compito di regressione e uno di classificazione. Risolvendo un problema di regressione, l'algoritmo NILM cerca di stimare il consumo di energia dei singoli apparecchi, mentre risolvendo un problema di classificazione, cerca di identificare quale apparecchio è in funzione in ogni stato.

A seconda dell'approccio NILM scelto, il task di classificazione consiste nell'identificare gli apparecchi o nel raggrupparli in cluster. Dall'articolo fondamentale di Hart [1], numerosi algoritmi di disaggregazione dell'energia sono stati proposti in letteratura per migliorare i suoi risultati [2, 3].

Gli approcci NILM possono essere divisi in metodi event-based o state-based. A seconda che la ground truth sia disponibile o meno, gli algoritmi NILM possono essere supervisionati o non supervisionati [4]. Gli approcci basati sugli eventi utilizzano tecniche di rilevamento dei bordi per catturare variazioni statisticamente significative del segnale. Gli approcci più comuni event-based coinvolgono un rilevamento non supervisionato degli eventi del segnale aggregato e una classificazione supervisionata per assegnare gli apparecchi noti agli eventi rilevati. Gli strumenti di classificazione che si possono trovare in letteratura sono basati su Support Vector Machines (SVM) [5], Decision Trees [6] e una combinazione di vari metodi di classificazione [7]. Inoltre, le tecniche di

clustering basate sul Dynamic Time Warping (DTW) sono utilizzate per identificare le finestre con modelli di consumo simili e per estrarre firme di carico rappresentative [8,9]. La ricerca recente sulla NILM basata sugli eventi coinvolge algoritmi di Graph Signal Processing (GSP) [10], sia supervisionati che non supervisionati, dove la correlazione spazio-temporale dei dati viene trovata incorporando i segnali in un grafico. Recentemente, sono stati suggeriti algoritmi di classificazione multi-label che impiegano caratteristiche del dominio temporale e del dominio wavelet (vedi [11] e riferimenti). Gli approcci basati sullo stato, invece, considerano ogni apparecchio come una macchina a stati finiti e disaggregano il segnale di potenza totale in base ai modelli di distribuzione del carico degli apparecchi. Al fine di esplorare tutte le possibili combinazioni tra le diverse sequenze di stato degli apparecchi, l'ottimizzazione combinatoria (CO) e i modelli di Markov nascosti (HMM) [12,13,14] sono comunemente usati tra gli approcci state-based. Tuttavia, quando il numero di dispositivi aumenta, il numero di combinazioni di stati aumenta esponenzialmente, aumentando la complessità del problema. Un altro limite di questi approcci è che tendono a produrre risultati insoddisfacenti in presenza di sommerso (carico non monitorato). In [15] si utilizzano HMM differenziali con l'algoritmo di massimizzazione delle aspettative per generare modelli di transizione di stato in modo non supervisionato. In [16] le misurazioni di sub-metering sono utilizzate per costruire HMM super-stato e l'inferenza viene eseguita attraverso un algoritmo sparso di Viterbi, ottenendo una stima efficiente del consumo di energia. In [17] si modella il problema della disaggregazione come un problema di fattorizzazioni di matrici. Tra gli approcci proposti di recente nella letteratura del NILM, è stata dedicata particolare attenzione alle Deep Neural Networks [18, 19, 20, 21].

2.1 Multi-task Deep Neural Network con meccanismo di Attention

L'architettura generale della rete riprende quella proposta da Piccialli and Sudoso [21]. Si tratta di un modello costituito da due ramificazioni interconnesse che si occupano separatamente di svolgere regressione e classificazione. Il ramo di regressione permette di determinare il consumo di un elettrodomestico in ogni istante di tempo, mentre il ramo di classificazione si occupa di determinare lo stato On/Off. Le due porzioni di rete non sono tra loro completamente separate: oltre a ricevere in ingresso il medesimo input, l'uscita prodotta in fase di classificazione concorre anche al perfezionamento di quella prodotta in fase di regressione. La regressione vera e propria viene svolta a valle dell'encoding dell'input, finalizzata a estrarre da esso le caratteristiche salienti del carico. Nell'encoder vengono impiegati layer convoluzionali e, dal momento che si lavora con serie storiche, anche layer ricorrenti. Questi ultimi sono in grado di sfruttare la nozione di tempo inclusa nei dati per migliorare la qualità delle feature estratte. Il ramo di regressione viene poi potenziato anche attraverso l'introduzione di un layer di attention, che ha come obiettivo quello di rafforzare la rappresentazione di porzioni di input ritenute particolarmente importanti al fine della disaggregazione. Lo schema generale della rete è riportato qui di seguito in Figura 9.

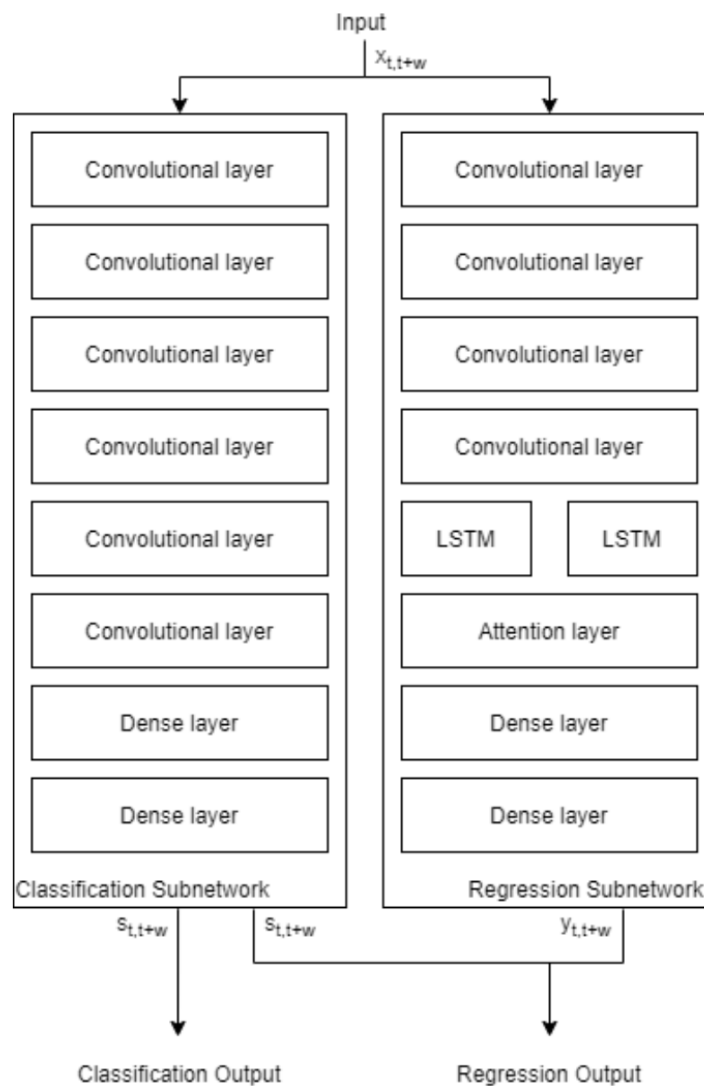


Figura 9 : Architettura della rete

La rete neurale prende in input finestre di dimensione L di valori $x_{t,L} = (x_t, \dots, x_{t+L})$ che, al variare di t , rappresentano il consumo energetico aggregato dei vari elettrodomestici nel tempo. In fase di addestramento, però, alla rete va fornito anche l'output corrispondente.

Nello specifico, per ogni finestra di input $x_{t,t+L}$, bisogna specificare due finestre di output $s_{t,L} = (s_t, \dots, s_{t+L})$ e $y_{t,L} = (y_t, \dots, y_{t+L})$.

Il vettore y rappresenta il consumo energetico dell'elettrodomestico, mentre il vettore s rappresenta lo stato dell'elettrodomestico nello stesso range temporale. In particolare, s è costituito unicamente da valori 0 e 1, dove 0 sta per OFF e 1 sta per ON. Tipicamente, gli smart meter permettono di ottenere direttamente il vettore y mentre il vettore s deve essere ricavato in modo indiretto [22]. A tale scopo definiamo un threshold in Watt per estrarre da y lo stato dell'elettrodomestico. In questo studio, il valore del threshold varia in un range che va dai 15 Watt ai 100 Watt, in base all'elettrodomestico da disaggregare.

Quando si lavora con reti neurali, è fondamentale normalizzare le finestre di input e di output. In questo caso è conveniente standardizzare con media zero e varianza unitaria i dati aggregati e

normalizzare con minimo e massimo i dati relativi al singolo elettrodomestico usati per le finestre di output.

Più in dettaglio, la sottorete di classificazione ha la seguente struttura, dove i parametri sono stati scelti in cross-validation:

1. Conv1D(filters=30, kernel size=10, strides=1, padding='valid', activation='relu')
2. Conv1D(filters=30, kernel size=8, strides=1, padding='valid', activation='relu')
3. Conv1D(filters=40, kernel size=6, strides=1, padding='valid', activation='relu')
4. Conv1D(filters=50, kernel size=5, strides=1, padding='valid', activation='relu')
5. Conv1D(filters=50, kernel size=5, strides=1, padding='valid', activation='relu')
6. Conv1D(filters=50, kernel size=5, strides=1, padding='valid', activation='relu')
7. Flatten()
8. Dense(units=1024, activation='relu')
9. Dense(units=windows size, activation='sigmoid')

La configurazione della sottorete di regressione è invece la seguente:

1. Conv1D(filters=16, kernel size=4, strides=1, padding='valid', activation='relu')
2. MaxPool1D(pool size=2)
3. Conv1D(filters=16, kernel size=4, strides=1, padding='valid', activation='relu')
4. MaxPool1D(pool size=2)
5. Conv1D(filters=16, kernel size=4, strides=1, padding='valid', activation='relu')
6. MaxPool1D(pool size=2)
7. Conv1D(filters=16, kernel size=4, strides=1, padding='valid', activation='relu')
8. MaxPool1D(pool size=2)
9. Bidirectional LSTM(units=256, activation='tanh', merge mode='concat')
10. Attention(units=256, activation='tanh')
11. Dense(units=256, activation='relu')
12. Dense(units=windows size, activation='relu')

La rete neurale deep è stata implementata da zero in Python ed addestrata utilizzando Keras, seguendo [21].

Keras è una libreria di alto livello che consente di comporre e connettere i layer delle reti neurali utilizzando il linguaggio Python. Keras è molto flessibile ed è in grado di interoperare con un certo numero di backend; tra questi, TensorFlow è quello utilizzato all'interno di questo studio. Il meccanismo attention nella rete di regressione è stato implementato con un layer custom tramite le primitive messe a disposizione da Tensorflow. Inoltre, TensorFlow è progettato per sfruttare la GPU al fine di velocizzare notevolmente il processo di addestramento. Per usufruire di questa feature, si è scelto di utilizzare Google Colab, cioè un ambiente di esecuzione online che fornisce allo sviluppatore una GPU NVIDIA.

La loss utilizzata per risolvere il problema della disaggregazione è la Binary CrossEntropy per l'output del ramo di classificazione e la Mean Squared Error per l'output della regressione. La loss complessiva del modello è calcolata semplicemente come la somma di questi due contributi. Il modello è stato addestrato utilizzando l'algoritmo Adam per minimizzare tale loss. Adam è una

variante della classica backpropagation a cui si aggiunge il momentum e la possibilità di utilizzare un learning rate ad hoc per ciascun parametro del modello. Al fine di migliorare le prestazioni di ogni modello, si è fatto tuning anche sui parametri dell'ottimizzatore. L'addestramento è stato configurato per durare al massimo 100 epoche; nella pratica, le epoche che vengono effettivamente completate sono inferiori a 100, perché interviene l'early stopping. L'early stopping è stato configurato per arrestare prematuramente l'addestramento quando, per un certo numero di epoche, la loss sul validation set smette di diminuire; in questo modo, si riesce a contenere il fenomeno dell'overfitting e a ottimizzare i tempi di addestramento. Infine, l'addestramento è avvenuto su una GPU NVIDIA TESLA T4 fornendo alla rete mini-batch costituiti da 32 campioni (sequenze di input ed output).

2.2 Risultati ottenuti tramite la rete deep

Per valutare le performance delle reti è necessario definire delle metriche. Queste sono divise in metriche energy-based e state-based; quest'ultime sono metriche di classificazione binaria e si focalizzano sulla rivelazione dello stato del dispositivo, in cui quello attuale e quello predetto sono stimati utilizzando threshold on/off specifici per ogni device. Si definiscono quindi le seguenti metriche:

1. $TP_i = \sum_{t=1}^T 1(s_i(t) = on, \hat{s}_i(t) = on)$ true positive del dispositivo i
2. $FP_i = \sum_{t=1}^T 1(s_i(t) = off, \hat{s}_i(t) = on)$ false positive del dispositivo i
3. $TN_i = \sum_{t=1}^T 1(s_i(t) = off, \hat{s}_i(t) = off)$ true negative del dispositivo i
4. $FN_i = \sum_{t=1}^T 1(s_i(t) = on, \hat{s}_i(t) = off)$ false negative del dispositivo i

dove con $1(x)$ si intende la funzione indicatrice che vale 1 se si verifica la condizione x , 0 altrimenti. Si indica con $s_i(t)$ lo stato effettivo del dispositivo i all'istante t , e con $\hat{s}_i(t)$ lo stato predetto del dispositivo i all'istante t . Definiamo anche la state-based precision PS_i e la state-based recall RS_i . La prima indica il rapporto di stati "on" correttamente classificati su tutti quelli etichettati come "on"; la seconda indica il rapporto di stati "on" correttamente classificati su tutti quelli che sarebbero da etichettare come "on":

$$PS_i = \frac{TP_i}{TP_i + FP_i}$$

$$RS_i = \frac{TP_i}{TP_i + FN_i}$$

A questo punto, è possibile introdurre la state-based F1-score FS_i come la media armonica tra precision e recall:

$$FS_i = \frac{2PS_i \cdot RS_i}{PS_i + RS_i}$$

Quanto migliore è la classificazione, quanto più questa metrica è vicina a 1. Le metriche energy-based valutano la bontà della regressione; non si basano su valori binari, ma sulla distanza che intercorre tra la curva dei consumi reali $y(t)$ e di quelli predetti $\widehat{y}(t)$. Definiamo la energy-based precision PE_i , come il rapporto tra tutta la potenza predetta e realmente consumata

dall'elettrodomestico, su tutta quella predetta; definiamo invece la energy-based recall RE_i , come il rapporto tra tutta la potenza predetta e realmente consumata dall'elettrodomestico, su tutta quella effettivamente consumata. Formalmente:

$$PE_i = \frac{\sum_{t=1}^T \min(y(t), \widehat{y}(t))}{\sum_{t=1}^T \widehat{y}(t)}$$

$$RE_i = \frac{\sum_{t=1}^T \min(y(t), \widehat{y}(t))}{\sum_{t=1}^T y(t)}$$

A questo punto, possiamo introdurre anche l'energy-based F1 score che è ancora la media armonica tra precision e recall, ma nel caso continuo:

$$FE_i = \frac{2PE_i \cdot RE_i}{PE_i + RE_i}$$

Anche in questo caso l'energy-based F1 score assume un valore in $[0,1]$: quanto migliore è la regressione ottenuta quanto più assume valori vicini a 1.

Riportiamo in Tabelle 8-10 i risultati ottenuti utilizzando la rete deep sui vari elettrodomestici nelle tabelle, utilizzando l'80% dei dati come training set e il 20% dei dati come test set.

Lavatrice	Casa 1	Casa 5	Casa 6	Casa 7	Casa 8
<i>FE</i>	0.74	0.53	0.74	0.8	0.47
<i>FS</i>	0.89	0.64	0.86	0.89	0.64

Tabella 8 : Energy F1 Score e State F1 Score per la lavatrice

Lavastoviglie	Casa 1	Casa 5	Casa 6	Casa 7	Casa 9
<i>FE</i>	0.83	0.51	0.75	0.74	0.92
<i>FS</i>	0.93	0.66	0.81	0.83	0.99

Tabella 9 : Energy F1 Score e State F1 Score per la lavastoviglie

Lavatrice	Casa 6	Casa 7	Casa 9
<i>FE</i>	0.85	0.81	0.72
<i>FS</i>	0.43	0.94	0.81

Tabella 10 : Energy F1 Score e State F1 Score per il frigorifero

I risultati ottenuti sono molto buoni, soprattutto considerando i dati grezzi che presentavano numerose criticità, e dati mancanti. A titolo di esempio, si riportano in Figura 10 i plot del consumo reale e del consumo predetto di alcuni elettrodomestici.

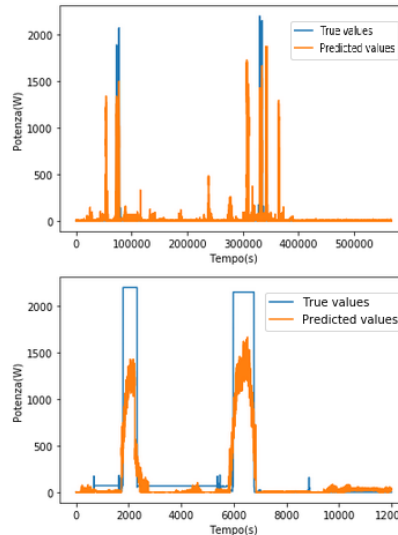


Figura 10: Lavastoviglie casa 5: il primo plot contiene l'intero test set, il secondo è uno zoom sulla prima attivazione

Notiamo inoltre come il meccanismo di attention sia effettivamente in grado di individuare gli istanti temporali corrispondenti ai cambi di stato dell'elettrodomestico. Riportiamo a questo scopo in Figura 11 i pesi di attention per la lavastoviglie di casa 9 in diversi istanti di tempo attraverso l'uso di un heatmap, dove un colore più scuro indica un valore maggiore. Come si può notare, per tutta la durata dell'attivazione, i pesi hanno una rilevanza maggiore permettendo così alla rete di dare maggiore importanza a questi timestamp nella fase di previsione.

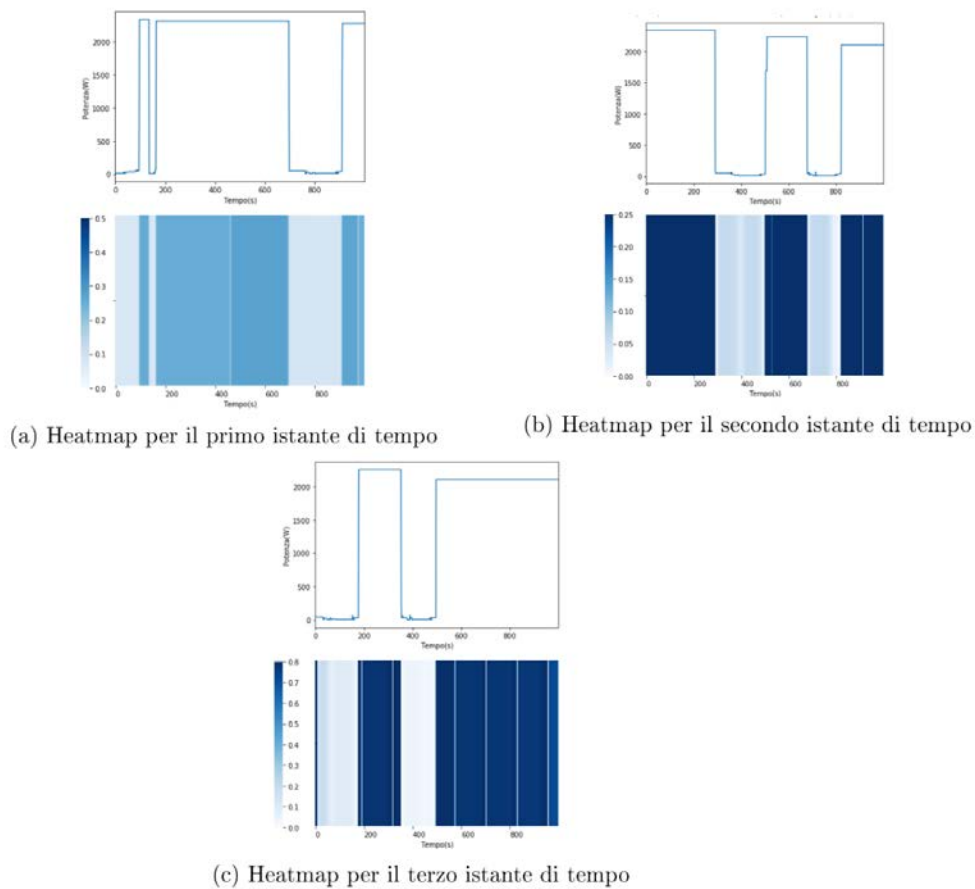


Figura 11: Heatmap dei pesi dell'attention per la lavastoviglie di casa 9

3 Clustering e previsione del consumo complessivo

Un aspetto interessante per lo studio e il miglioramento dei comportamenti energetici, è quello di saper prevedere i consumi complessivi di una certa abitazione con una particolare attenzione ai picchi. Questi vanno infatti evitati per minimizzare la produzione di energia (e quindi l'utilizzo di combustibili fossili), evitare blackout sulla rete e diminuire i costi. Dal punto di vista scientifico questo problema corrisponde alla previsione di una serie storica univariata che rappresenta la misurazione dell'energy meter in una sequenza di istanti temporali. La difficoltà consiste nell'irregolarità della serie da prevedere, che è influenzata da diversi fattori tra cui il clima, le caratteristiche dell'immobile considerato (locazione geografica, il sistema di illuminazione/ventilazione installato, i materiali con cui è stato realizzato), le caratteristiche temporali (giorno e fascia oraria della rilevazione) e soprattutto il comportamento degli occupanti. Quest'ultimo è indipendente dai consumi e non è "costante" nei giorni. Nel corso del tempo, all'interno di una stessa fascia oraria, ogni utente può mostrare una variazione sia nella quantità dei dispositivi utilizzati che nel loro tempo di utilizzo, arrivando ad avere comportamenti energetici profondamente differenti. Tale caratteristica rende la varianza dei consumi molto elevata

soprattutto nel caso degli edifici residenziali, aumentando fortemente la difficoltà nella previsione di quelli futuri. Tipicamente per migliorare la previsione vengono introdotte delle informazioni utili a caratterizzare, almeno in parte, il comportamento degli utenti. Questo introduce una seconda complessità al problema. Un'analisi su scala piccola può essere realizzata facilmente installando diversi sensori nell'immobile considerato per ricavare delle informazioni di supporto. In questo caso si parla di *sensor-based analysis*. Replicare l'analisi su larga scala comporta una spesa molto elevata da sostenere, sia a livello economico che tecnico vista la mole di sensori da installare. Inoltre non vanno sottovalutati aspetti relativi alla *privacy* che potrebbe impedire l'installazione di diversi sensori. In questo lavoro, ci si è basato solo sui consumi storici, integrati con dati meteorologici, semplicemente reperibili tramite servizi online, ed informazioni socio-demografiche ricavate dai questionari compilati dagli occupanti delle diverse abitazioni. L'utilizzo dei questionari rappresenta infatti un'alternativa utile per ricavare preziose informazioni sul comportamento degli utenti evitando limitazioni tecniche, economiche e legate alla *privacy* nell'uso di eventuali sensori.

L'eterogeneità che caratterizza i consumi elettrici porta al bisogno di profilare gli utenti in base al loro comportamento energetico. In questo modo è possibile definire un modello di previsione che può essere applicato su un set di utenti caratterizzati da un "profilo energetico" simile. Tale profilazione può essere definita applicando tecniche di *clustering*. Definire dei cluster di utenti è un vantaggio significativo per lo studio dei loro consumi, e può essere applicato per diversi scopi che vanno dalla previsione dei consumi futuri, alla definizione di un *pricing* specifico per ogni profilo identificato. Analogamente al problema della previsione, la maggiore difficoltà è rappresentata dalle abitazioni residenziali a causa della loro varianza nei consumi. Diversi studi hanno dimostrato che una clusterizzazione di qualità non può essere definita esclusivamente sulla base dei consumi elettrici. Anche in questo caso è necessario integrare informazioni di supporto utili per identificare pattern comportamentali che difficilmente possono essere individuati dai soli consumi.

Nel nostro caso abbiamo deciso di definire cluster di utenti, in modo da avere una visione di insieme sui comportamenti dei diversi residenti, e abbiamo anche tentato la strada di addestrare un modello di previsione per ogni cluster, da utilizzare poi per prevedere i consumi futuri di tutti gli utenti appartenenti al medesimo cluster (scenario *unseen*). La possibilità di addestrare il modello su un singolo utente ma di poterlo utilizzare per ricavare delle previsioni su tutti gli utenti del medesimo cluster, rende il modello facilmente replicabile su larga scala. Chiaramente, la bontà delle previsioni ricavate dipende dalla validità dei profili energetici individuati, e dalla ricchezza delle informazioni utilizzate per costruire sia i cluster che le previsioni.

3.1 Letteratura correlata

In letteratura, diverse tecniche di previsione di consumi elettrici sono state proposte. L.M. Candanedo et al. in [25] hanno studiato l'accuratezza nella previsione dei modelli SVM, Random Forest e Gradient Boosting. Lo studio è basato sui consumi di un'abitazione residenziale, collezionati ad intervalli di 10 minuti, a cui sono state integrate informazioni relative alla temperatura e all'umidità dell'abitazione rilevate attraverso appositi sensori. Lo studio dimostra la bontà dei risultati ottenuti con il Gradient Boosting rispetto agli altri modelli testati. Nell'area degli immobili residenziali diversi studi hanno testato i vantaggi nell'integrazione di informazioni di carattere

demografico, relative all'immobile considerato, agli elettrodomestici utilizzati e all'occupazione degli utenti. R. Edwards e J. New [26] hanno analizzato i consumi di tre abitazioni in cui sono stati installati circa 140 sensori per collezionare informazioni sul comportamento umano tra cui pattern sull'utilizzo degli elettrodomestici e sull'occupazione dell'immobile. Tale studio, a causa della numerosità delle informazioni da utilizzare, risulta difficilmente replicabile su larga scala a meno di un importante investimento economico. Z. Wang e Y. Ding [27] hanno sviluppato invece un modello per studiare la relazione tra i consumi e il comportamento delle persone, valutando in particolare la sua variazione rispetto a diversi meccanismi di pricing applicati. L'obiettivo del lavoro presentato è quello di stimare l'energia potenzialmente risparmiata in un ufficio utilizzato da 20 persone. Lo studio conferma che all'aumento del pricing applicato corrisponde un comportamento più "green" dei dipendenti. Tale strategia è infatti applicata dal governo Cinese il quale ha introdotto meccanismi di pricing relativi all'andamento dei consumi (critical peak pricing, peak-valley pricing) con l'obiettivo di incentivare i cittadini ad un uso efficiente dei dispositivi elettrici. Per ridurre l'impatto casuale del comportamento di ogni utente sui consumi di una singola abitazione, alcuni studi hanno concentrato l'attenzione sui consumi degli edifici residenziali o commerciali. In particolare Rishee K. Jain et al. [28] hanno presentato un modello di previsione per residenze plurifamiliari in cui i consumi delle singole abitazioni sono stati aggregati. La previsione è stata ricavata utilizzando una SVM con consumi di diversa granularità, dimostrando che i risultati migliori sono stati ottenuti utilizzando i consumi orari. K. Grolinger et al. [29] hanno studiato invece i consumi di un impianto per grandi eventi in Ontario, Canada. In questo caso sono stati utilizzati una SVM, una rete neurale e consumi con tre granularità: giornaliera, oraria e di 15 minuti. Inoltre, per migliorare l'accuratezza della previsione, sono state integrate informazioni di carattere meteorologico e temporale. Lo studio mostra che i risultati migliori sono ottenuti dalla rete neurale con consumi giornalieri. L'analisi degli impianti per eventi è importante a causa dei consumi e dei relativi costi molto elevati. Utilizzare l'energia in modo appropriato in questi contesti comporterebbe infatti notevoli benefici. Un ulteriore aspetto del problema riguarda la granularità dei consumi. In generale la scelta ottimale, cioè che riesce a catturare maggiormente la varianza dei consumi, dipende dal tipo di immobile considerato e dall'obiettivo stesso dell'analisi. Per le abitazioni residenziali, nel caso in cui si vogliono offrire informazioni in tempo reale all'utente, in modo da informarlo sui consumi ed evitare possibili picchi, le previsioni avvengono su consumi aggregati su intervalli temporali solitamente entro l'ora. Nel caso di analisi legate alla pianificazione della distribuzione dell'energia nella rete elettrica o alla definizione del pricing si possono invece ipotizzare consumi di granularità maggiore. R. K. Jain et al. [30] hanno studiato l'impatto della diversa granularità nella previsione dei consumi di un edificio residenziale. Lo studio valuta le previsioni ricavate da una SVM utilizzando consumi con granularità giornaliera, oraria e di 10 minuti. In questo caso la granularità che riesce a catturare maggiormente la varianza dei consumi è quella oraria. Una particolarità dello studio è il fatto che viene considerata anche la granularità "spaziale" dei consumi, testando l'utilizzo dei consumi dell'intero edificio oppure dividendoli per unità o per piano dell'immobile. Lo studio dimostra che quest'ultimi permettono di ottenere i risultati migliori. Ultimamente, sono molto in voga strumenti basati sul deep learning, che prevedono l'uso di reti neurali molto profonde, con un costo computazionale molto alto e che richiedono una mole di dati molto elevata. A. Almalaq e J.J. Zhang [31] hanno proposto un modello ibrido basato su un algoritmo genetico ed una rete neurale LSTM, con l'obiettivo di prevedere i consumi di edifici residenziali e commerciali. Nello specifico i parametri di ogni layer della rete neurale (ad esempio il numero di neuroni) vengono definiti in modo dinamico dall'algoritmo genetico. L'approccio presentato è caratterizzato da un elevato costo computazionale che rende necessaria una sua esecuzione in ambiente distribuito. I risultati dimostrano comunque l'efficacia nell'utilizzo di modelli più

complessi. Anche la profilazione energetica dei clienti è oggetto di molti studi che hanno esplorato i diversi aspetti che caratterizzano il problema.

F. Iglesias e W. Kastner [32] hanno studiato diverse combinazioni di algoritmi di clustering e distanze per valutare quella in grado di stabilire la profilazione migliore. A tale scopo sono state definite diverse metriche per valutare la similitudine tra gli elementi di ogni cluster. Lo studio è basato sui consumi orari di diversi edifici universitari in Spagna, registrati tra il 2011 e il 2012.

J. du Toit et al. [33] hanno testato i risultati dell'algoritmo di k-means utilizzando diverse metriche. In particolare hanno studiato le differenze tra le i cluster ottenuti sulla base dei consumi di diversi immobili industriali registrati dal 2007 e il 2014. Nell'analisi proposta il dataset è stato ridotto attraverso la tecnica delle componenti principali (PCA) in modo tale da testare gli effetti di quest'ultima. I risultati dimostrano l'utilità di tale tecnica, la cui applicazione non comporta effetti significativi sui cluster ottenuti. Un aspetto importante affrontato da diversi studi è rappresentato dalla relazione che lega il comportamento degli utenti ai picchi di consumo. A. Satre-Meloy et al. [34] hanno integrato ai consumi di diverse abitazioni londinesi i dati relativi all'attività degli utenti per profilare il comportamento e prevederne i consumi futuri. Lo studio dimostra che introducendo informazioni relative alle abitudini degli utenti si può migliorare la loro profilazione, e si è in grado di individuare dei pattern correlati ai picchi registrati nei consumi.

3.2 Generazione dei cluster relativi ai consumi elettrici

Per determinare i profili di consumi energetici tramite algoritmo di clustering, abbiamo considerato seguendo la letteratura, sia features derivate dai dati storici dei consumi elettrici (lo smart meter) che features relative al comportamento energetico degli abitanti e al funzionamento dei diversi elettrodomestici, estratte dai questionari compilati dagli abitanti delle diverse abitazioni. I questionari sono stati compilati dai proprietari delle abitazioni e contengono informazioni di diversa natura. In una prima parte ci sono dati anagrafici dell'utente e informazioni di natura geografica sull'abitazione in oggetto. Seguono caratteristiche architettoniche come il numero di piani e la metratura, quindi dettagli riguardo l'impianto di riscaldamento e di raffreddamento oltre al numero di occupanti nelle diverse fasce orarie della giornata.

Di seguito viene riportato l'elenco delle informazioni principali contenute nel questionario, organizzate per tipologia:

1. Anagrafica utente;
2. Numero di occupanti per fascia oraria;
3. Caratteristiche architettoniche;
 - I. Numero di piani
 - II. Numero di stanze
 - III. Numero di finestre
4. Dati geometrici sull'abitazione:
 - I. Superficie calpestabile (totale, per piano)
 - II. Superficie lorda (totale, per piano)
 - III. Volume totale (netto, lordo)
5. Elettrodomestici della cucina:

- I. Tipo (gas, elettrico)
 - II. Uso giornaliero (minuti)
6. Dispositivi audio-video;
- I. Quantità
 - II. Classe energetica
 - III. Uso giornaliero (minuti)
7. Illuminazione;
- I. Numero di lampade obsolete
 - II. Numero di lampade tradizionali
 - III. Numero di lampade a risparmio energetico
 - IV. Numero di lampade ad alta efficienza
8. Consumi e costi energetici;
- I. Energia elettrica
 - II. Gas naturale
 - III. Costi

I dati ritenuti utili per la caratterizzazione delle abitazioni e delle abitudini degli occupanti, in relazione ai consumi elettrici, sono relativi a:

1. Numero di occupanti nelle fasce orarie “mattina”, “pomeriggio”, “sera” e “notte”;
2. Classe energetica, quantità e utilizzo medio degli elettrodomestici.

La classe energetica è un attributo categorico ordinale, che è utilizzata per esprimere l’efficienza energetica di un elettrodomestico attraverso il suo consumo elettrico in kWh annui. Per convenzione viene rappresentata mediante una lettera tra la A e la D, dove A+++ e D rappresentano rispettivamente la classe energetica migliore e peggiore. Ogni classe energetica è stata sostituita con il relativo consumo in kWh. L’associazione tra la classe energetica ed il relativo consumo in kWh è stata definita facendo riferimento al documento Enea [35]. Ad ogni classe energetica non corrisponde un valore puntuale del consumo ma un intervallo, cioè un range del possibile consumo. Presi due elettrodomestici aventi stessa classe energetica, il loro consumo energetico non sarà perfettamente uguale ma varierà, seppur in maniera limitata, in base alla casa produttrice, al tipo specifico di elettrodomestico (una versione più o meno recente del dispositivo) e alla qualità dei componenti utilizzati nella produzione. Per questo motivo la classe energetica è stata sostituita dal valore medio del relativo intervallo di consumo. Nel caso in cui un elettrodomestico non risulta essere presente nell’abitazione ed è disponibile l’informazione relativa alla sua classe energetica, quest’ultima è stata sostituita con il valore zero. Tale scelta non ha introdotto un rumore nel dataset in quanto, l’assenza di un elettrodomestico, può essere rappresentata correttamente attraverso un suo consumo nullo. Come features derivate dai dati storici dei consumi elettrici, per ogni abitazione sono stati calcolati il valore medio e la deviazione standard del consumo relativi alle fasce orarie 00:00 - 08:00, 08:00 - 16:00 e 16:00 -00:00.

Per procedere alla clusterizzazione sono stati selezionati gli attributi utili per caratterizzare il consumo elettrico. Il primo passo è stato l’eliminazione di tutti quelli che non hanno relazione con il consumo elettrico (ad esempio gli attributi relativi al consumo di gas) o caratterizzati da un valore

costante per tutte le abitazioni (come, ad esempio, la presenza di una connessione internet di tipo "flat"). Questi ultimi, infatti, non sono utili per caratterizzare il consumo elettrico di un'abitazione rispetto agli altri.

Per avere un focus sul consumo effettivo degli elettrodomestici si è stimato il loro consumo "effettivo" come prodotto tra il tempo medio di utilizzo del dispositivo e il relativo consumo medio (che, come detto, è stato stimato in base alla classe energetica). Nei casi in cui non si dispone di entrambe le informazioni si è utilizzata la sola a disposizione in modo da caratterizzare, come possibile, l'utilizzo di ogni elettrodomestico. Un esempio è rappresentato dall'illuminazione, caratterizzata dal tipo e dalla quantità di lampade installate non avendo informazioni relative al tempo medio di utilizzo.

Un ulteriore aspetto che è stato gestito riguarda la presenza in un'abitazione di più elettrodomestici dello stesso tipo. È il caso, per esempio, della tv che nelle abitazioni analizzate è presente in un numero di unità che varia tra zero e quattro. In questi casi si è scelto di aggregare i consumi medi individuali per ricavarne un consumo medio totale. In questo modo si è cercato di evitare di caratterizzare le abitazioni in base alla sola quantità di elettrodomestici posseduti, cercando invece una caratterizzazione basata sul loro utilizzo.

Infine, si è scelto di valutare per la clusterizzazione soltanto gli elettrodomestici considerati più significativi ai fini del consumo elettrico totale. Si è evitato quindi di considerare quelli utilizzati saltuariamente o caratterizzati da un tempo medio di utilizzo di qualche minuto. Tale scelta ha portato ad escludere, per esempio nel caso della cucina, il frullatore e la macchina del caffè in favore del forno elettrico ed a microonde.

Le scelte fatte per ogni categoria di consumo sono riassunte di seguito.

Illuminazione: dal questionario si è ricavata la quantità di lampade installate nelle abitazioni per ogni categoria (obsolete, tradizionali, risparmio energetico e alta efficienza). Si è scelto di non inserire il totale delle lampade installate per dare un peso maggiore alla loro tipologia. L'obiettivo è di evitare la creazione di cluster di abitazioni caratterizzate da una quantità simile di lampade installate ma di tipologia molto differente. Il documento Enea [35] conferma che è opportuno considerare l'illuminazione in quanto è stato stimato che contribuisce a circa l'11% del consumo elettrico totale di un'abitazione.

Raffreddamento: per i climatizzatori si è considerata la classe energetica e il numero di stanze climatizzate. Per considerare il potenziale consumo dei climatizzatori è stato calcolato il prodotto tra il consumo medio dei climatizzatori installati, ricavato dalla classe energetica, e il numero di stanze climatizzate.

Refrigerazione: sono stati considerati il frigorifero ed il congelatore per i quali è stata considerata la classe energetica. Entrambi hanno un forte impatto sul consumo elettrico totale dal momento che sono caratterizzati da un'alimentazione continua.

Elettrodomestici audio-video e informatici: tra gli elettrodomestici di tipo audio-video sono stati considerati monitor e tv essendo caratterizzati da un tempo medio di utilizzo quotidiano significativo. Come anticipato, vista la possibile presenza di più unità di tali dispositivi, i consumi medi individuali sono state aggregati ottenendo un consumo medio totale dell'elettrodomestico.

Non sono stati considerati dispositivi quali decoder, lettore dvd, radio e hi-fi in quanto caratterizzati da un utilizzo saltuario. Dal punto di vista informatico si sono tenuti in considerazione i consumi medi di pc fissi e portatili, ignorando invece le informazioni relative alle stampanti installate.

Cucina: non sono stati considerati elettrodomestici come frullatore, macchina del caffè, tostapane e robot perché caratterizzati da un tempo medio di utilizzo, e quindi un consumo elettrico, molto contenuto. Si è preferito invece considerare il forno elettrico e il forno a microonde, entrambi caratterizzati da un tempo medio di utilizzo quotidiano elevato e continuo durante tutto l'anno.

Cura della persona: è stato considerato il tempo medio di utilizzo dell'asciugacapelli in quanto molto dispendioso dal punto di vista elettrico, evitando invece di inserire la piastra per capelli. Dai questionari risulta infatti essere utilizzata per pochi minuti al giorno da una sola famiglia.

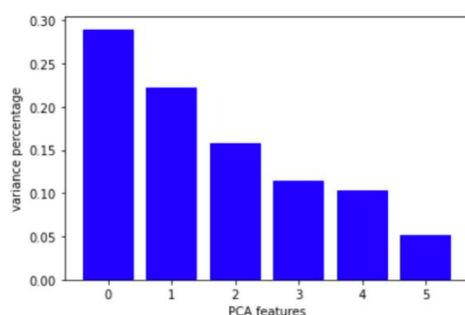
Lavaggio e asciugatura: sono state considerate lavatrici, asciugatrici, lavasciuga, aspirapolvere e ferro da stiro (sia con la caldaia che senza). Analogamente alla tv anche i consumi dei due tipi di ferro da stiro sono stati aggregati. Nel caso specifico di lavatrice, asciugatrice e lavasciuga, avendo a disposizione il numero di cicli settimanali è stato stimato il consumo medio settimanale ipotizzando una durata media di due ore per ogni ciclo.

In definitiva le features considerate per il clustering relative sia ai consumi dei singoli elettrodomestici che ai dati storici del meter sono riassunte in Tabella 11:

Consumo abitazione	Tempo di utilizzo	Consumo	Classe energetica	Quantità
Media 00:00-08:00	Forno	Tv	Frigorifero	Lampade
Media 08:00-16:00	Microonde	Climatizzatore	Congelatore	Pc portatile
Media 16:00-00:00	Aspirapolvere	Asciugatrice		
DevSt 00:00-08:00	Ferro da stiro	Lavatrice		
DevSt 08:00-16:00	Asciugacapelli	Lavastoviglie		
DevSt 16:00-00:00	Pc fisso	Condizionatore		

Tabella 11: Features utilizzate per il clustering

Vista la dimensione ridotta del dataset, si è applicata la Principal Component Analysis per ridurre il numero delle features. Il primo passo per applicare la PCA è la standardizzazione delle variabili quantitative in modo da ricondurle ad avere una distribuzione standard, cioè con media nulla e varianza pari ad uno. Attraverso la standardizzazione le variabili vengono ricondotte ad una stessa scala, rendendo possibile il confronto tra variabili caratterizzate da distribuzioni o scale diverse. Il secondo passo prevede lo studio della correlazione tra le variabili attraverso la definizione della matrice delle covarianze, costruita calcolando la correlazione tra le diverse coppie di variabili. La matrice delle covarianze viene utilizzata per costruire le componenti principali, cioè le nuove variabili restituite dalla PCA calcolate come combinazioni lineari pesate degli attributi iniziali e caratterizzate dal fatto di essere non-correlate tra loro. Da un punto di vista geometrico, le componenti principali possono essere interpretate come le direzioni verso cui i dati mostrano il massimo ammontare di varianza. Questa, di fatto, può essere messa in relazione con la quantità di informazione contenuta nei dati. Precisamente, possiamo dire che maggiore è la varianza dei dati (graficamente rappresentata dalla dispersione dei punti) e maggiore è la quantità di informazione che essi contengono. Le componenti principali vengono quindi scelte come le direzioni lungo cui i punti hanno la dispersione maggiore. Geometricamente, l'assenza di correlazione tra le componenti implica invece che le direzioni sono perpendicolari tra loro. Ogni componente viene costruita definendo, dalla matrice delle covarianze, un autovettore ed il relativo autovalore. Questi rappresentano rispettivamente la direzione verso cui c'è la varianza maggiore e l'ammontare di varianza "contenuta" nella componente. Gli autovettori vengono ordinati per autovalori decrescenti in modo tale da ottenere le componenti principali in ordine di importanza, cioè di informazione contenuta. Costruire le componenti in questo modo permette di comprimere in esse l'aspetto informativo delle variabili di partenza in misura via via decrescente. Ciò è fondamentale perché offre la possibilità di selezionare il numero minimo di componenti principali contenenti una percentuale fissata della varianza totale, evitando l'uso di tutte le altre componenti. Nel caso analizzato si è scelto di utilizzare sei componenti principali contenenti il 90% della varianza del dataset iniziale, come illustrato in Figura 12.



Componente	Percentuale di Varianza
0	0.2898
1	0.2227
2	0.1572
3	0.1146
4	0.1030
5	0.0521

Figura 12: Componenti principali e relativa varianza spiegata

Le componenti principali sono calcolate attraverso una combinazione lineare pesata degli attributi che caratterizzano il dataset utilizzato. Sono stati quindi ricavati gli attributi che hanno contribuito maggiormente, cioè quelli caratterizzati dai pesi maggiori nelle combinazioni. In ordine di importanza sono:

1. consumo del congelatore;
2. consumo della tv;
3. quantità di lampade di tipo tradizionale;
4. quantità di lampade di tipo risparmio energetico;
5. consumo dell'asciugatrice;
6. consumo del frigorifero.

Gli attributi nell'elenco sono in linea con le ipotesi fatte durante l'analisi dei questionari. Le differenze significative tra i profili energetici delle abitazioni vanno ricercate nel tempo medio di utilizzo e nella classe energetica degli elettrodomestici elencati. Come detto l'illuminazione ha un forte impatto sui consumi totali (11% circa) il quale può variare sia in base alla quantità che al tipo di lampade installate. Il frigorifero e il congelatore sono elettrodomestici caratterizzati da un'alimentazione continua, per cui diventa fondamentale considerarne la classe energetica. Per quanto riguarda la televisione si può ipotizzare che la differenza, oltre che nella quantità di dispositivi installati, è legata alle abitudini degli occupanti. L'informazione più interessante riguarda invece la presenza dell'asciugatrice. In questo caso si può ipotizzare che la differenza è legata alla presenza o meno dell'elettrodomestico, quindi al suo contributo al consumo totale. Un'osservazione importante riguarda l'assenza delle variabili legate al consumo rilevato dal meter, cioè il valore medio e la deviazione standard dell'energia nelle diverse fasce orarie. Questo significa che nella clusterizzazione ottenuta tali elementi non riescono a spiegare in modo soddisfacente la varianza che caratterizza le abitazioni considerate. Quest'ultima è legata a diversi aspetti tra cui il comportamento degli occupanti, fattori temporali e climatici. Ricavare eventuali pattern di comportamento energetico esclusivamente dal consumo totale è estremamente difficile, per cui diventa necessario integrare informazioni di supporto utili a caratterizzare le abitudini dell'utente. Costruito il dataset tramite PCA, si è scelto di effettuare il clustering tramite algoritmo di k-means, che è un'euristica utilizzata largamente per il clustering, e che usa come misura di distanza tra due punti del dataset la distanza euclidea al quadrato.

L'algoritmo di k-means [36] risolve il seguente problema di ottimizzazione:

$$\begin{aligned} \min_{\delta, z} \sum_{i=1}^N \sum_{j=1}^M \delta_{ij} \frac{1}{2} \|x_i - z_j\|^2 \\ \sum_{j=1}^M \delta_{ij} = 1 \quad i = 1, \dots, N \quad (\text{MSSC}) \\ \delta_{ij} \geq 0 \quad i = 1, \dots, N, j = 1 \dots M, z_j \in \mathbb{R}^p \end{aligned}$$

con x_i l'elemento i-esimo del dataset, z_j il centroide j-esimo e δ_{ij} variabile binaria che vale uno o zero rispettivamente se l'elemento i-esimo appartiene o meno al cluster j-esimo. L'algoritmo di k-means è composto da due fasi: si fissano i centroidi z_j e si calcola l'assegnamento ottimo δ_{ij} (che si può calcolare in forma chiusa e corrisponde ad assegnare ogni punto al centroide più vicino); dato l'assegnamento corrente si scelgono i centroidi ottimi (che corrispondono al baricentro dei punti appartenenti a ogni cluster).

Le due fasi vengono iterate fino a che i centroidi e l'assegnamento non cambiano più (e questo avviene in un numero finito di passi). Il punto restituito è un ottimo locale del problema (MSSC), e

la sua qualità dipende dall'inizializzazione dei centroidi iniziale. Per questo motivo, tipicamente l'algoritmo viene usato lanciandolo da un certo numero di inizializzazioni random diverse. Tra le limitazioni che caratterizzano l'algoritmo di k-means, e in generale tutti gli algoritmi di tipo "combinatorio", c'è il fatto di dover fissare il numero di cluster che l'algoritmo dovrà restituire. Esistono diversi metodi che possono essere utilizzati come supporto a tale scelta. Uno dei metodi più popolari è costituito dall'elbow method. L'elbow method è un metodo euristico per determinare il numero "ottimo" di cluster da creare, basato sulla minimizzazione della distanza intra-cluster. Il metodo prevede di eseguire più volte la clusterizzazione del dataset aumentando, ad ogni iterazione, il numero di cluster definiti. Per ogni clusterizzazione viene calcolata l'inertia cioè la somma delle distanze al quadrato tra i punti ed il centroide più vicino. I valori di inertia ottenuti vengono rappresentati graficamente con l'obiettivo di individuare il kink, o gomito, della curva ottenuta (caratteristica che motiva il nome stesso del metodo). Il numero minimo di cluster suggerito dal metodo è quello in corrispondenza del kink.

Si può notare che l'inertia, e quindi la distanza intra-cluster, diminuisce all'aumentare del numero di cluster restituiti. Tale aspetto impedisce di cercare la soluzione in corrispondenza del valore minimo di inertia dal momento che questo si ha sempre nel caso limite, cioè quando il numero di cluster restituiti è pari a quello degli elementi da clusterizzare. Attraverso il kink si va invece alla ricerca della diminuzione più significativa dell'inertia in modo tale da ottenere, da una parte, una distanza intra-cluster sufficientemente piccola e, dall'altra, un numero di cluster contenuto. Da un punto di vista computazionale si preferisce avere meno cluster con una cardinalità significativa che averne una quantità maggiore di cardinalità estremamente ridotte. L'andamento dell'inertia ottenuto per il dataset considerato è rappresentato in Figura 13.

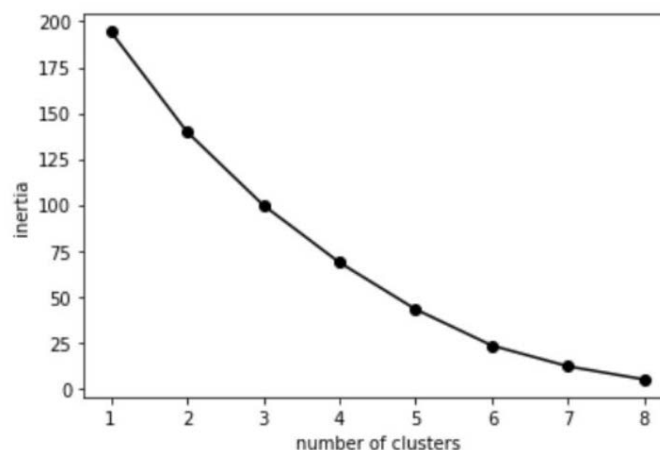
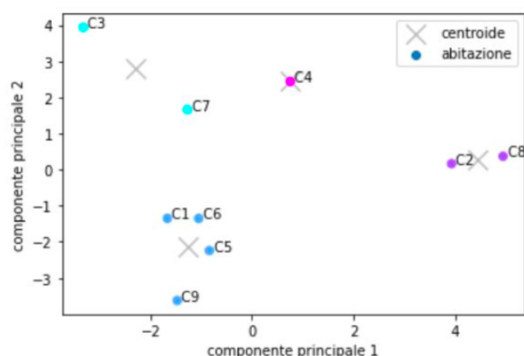


Figura 13: inertia nel caso in esame

In questo caso l'elbow method suggerisce un numero minimo di cluster intorno a cinque. Si può osservare infatti che la diminuzione dell'inertia inizia ad "addolcirsi" per valori maggiori di cinque. Nel caso analizzato si è scelto di utilizzare quattro cluster in quanto, osservando le diverse clusterizzazioni, questa è risultata essere la cardinalità minima per ottenere un cluster "ad-hoc" per l'abitazione C4, che è l'unica abitata da un solo residente. La clusterizzazione ottenuta dall'algoritmo di k-means, fissando il numero di cluster a quattro, è rappresentata in Figura 14. In quest'ultima la composizione di ogni cluster riporta le abitazioni dalla più vicina al centroide alla più lontana. La

rappresentazione grafica è stata realizzata rispetto alle prime due componenti principali restituite dalla PCA essendo quelle più rappresentative del dataset.



Cluster	Abitazioni
1	C7, C3
2	C1, C5, C6, C9
3	C4
4	C2, C8

Figura 14: cluster prodotti da kmeans

Ottenuti i cluster si è cercata una loro interpretazione basata sull’osservazione degli attributi con maggiore impatto nella definizione delle componenti principali. Come detto sono il consumo medio di congelatore e tv, quantità utilizzata di lampade tradizionali ed a risparmio energetico, consumo medio di asciugatrice e frigorifero. Di seguito sono elencate le evidenze principali.

Il cluster “quattro”, composto dalle abitazioni C2 e C8, mostra per gli elettrodomestici dei consumi medi per lo più nulli o comunque contenuti. L’illuminazione è invece caratterizzata dall’uso esclusivo di lampade a risparmio energetico. Si può concludere che il cluster comprende abitazioni caratterizzate da un profilo energetico contenuto.

Il cluster “due” composto dalle abitazioni C1, C5, C6 e C9, mostra un comportamento più “bilanciato” con consumi degli elettrodomestici mediamente più elevati e l’utilizzo di lampade di entrambe le tipologie. Dunque, il cluster comprende abitazioni caratterizzate da un consumo più elevato rispetto al precedente.

L’abitazione C4, l’unica ad essere abitata da un unico soggetto, è stata isolata nel cluster “tre”. Come detto l’aspettativa è che sia l’abitazione caratterizzata dal consumo totale minore. L’ipotesi risulta vera per quanto riguarda i consumi medi degli elettrodomestici mentre l’illuminazione ha un peso sui consumi maggiore del previsto. La causa è dovuta alla tipologia di lampade installate, essendo l’abitazione C4 quella caratterizzata dalla quantità maggiore di lampade tradizionali.

Il cluster "uno", composto dalle abitazioni C7 e C3, mostra i consumi medi degli elettrodomestici che sono, come prevedibile, maggiori rispetto all’abitazione C4 mentre l’illuminazione vede un utilizzo maggiore di lampade a risparmio energetico. Questo aspetto rende il profilo energetico del cluster più "virtuoso" rispetto alla C4, ed è un risultato inatteso dal momento che le abitazioni che lo compongono sono abitate da un numero di individui che varia da tre a quattro. Per motivare tale aspetto i consumi delle abitazioni sono stati confrontati in base alle fasce orarie. Dal confronto è emerso che le abitazioni C3 e C7 sono caratterizzate da un consumo minore, rispetto alla C4, nelle fasce orarie comprese nell’intervallo 00:00 – 16:00. Si può ipotizzare che, nonostante il numero di utenti, la differenza nei consumi elettrici è strettamente legata alle abitudini quotidiane dei soggetti. L’utente dell’abitazione C4 mostra infatti un maggiore utilizzo di tv, pc e forno elettrico mentre gli utenti dell’abitazione C3, oltre ad un utilizzo più contenuto di tv e pc, sono gli unici ad essere caratterizzati dall’utilizzo di un forno a gas. La C7 risulta invece bilanciare i consumi maggiori della tv con l’utilizzo di elettrodomestici di classe energetica superiore e il maggiore utilizzo di lampade a risparmio energetico rispetto alle altre due.

3.3 Previsione dei consumi elettrici

Una volta ottenuti i cluster dei consumi energetici, si è passata alla fase successiva del lavoro, ovvero la previsione dei consumi elettrici complessivi delle diverse abitazioni. Il primo passo è stata la costruzione del dataset, in cui sono stati inclusi dati storici sui consumi complessivi combinati con dati meteorologici, come suggerito in letteratura. Infatti, diversi studi basati sulla previsione dei consumi elettrici hanno dimostrato l'effettivo miglioramento dovuto all'inserimento di dati di carattere meteorologico. Ad esempio T. Wijaya et al. [37] hanno confrontato l'accuratezza di una previsione di tipo aggregato rispetto alle previsioni individuali ricavate dai modelli Linear Regression, Multi-Layer Perceptron e Support Vector Regression. Lo studio è basato sui consumi di oltre 5000 edifici residenziali irlandesi, rilevati tra il 2009 e il 2010 ad intervalli di 30 minuti. Per migliorare l'accuratezza delle previsioni sono state utilizzate anche informazioni di tipo meteorologico e temporale. I risultati, oltre a mostrare i benefici della previsione di tipo aggregato, confermano la loro utilità a supporto dei consumi. Il ruolo importante che giocano ai fini della previsione è dovuto alla stretta correlazione del meteo con gli stessi consumi elettrici. I dati utilizzati per l'analisi presentata sono stati ricavati da World Weather Online [38], un servizio online che offre un database contenente informazioni meteorologiche per diverse città nel mondo.

Il dataset utilizzato per la fase di previsione è stato definito aggiungendo al consumo elettrico le informazioni di natura climatica: temperatura media, pressione e umidità. Inoltre, sono state aggiunte informazioni di natura temporale relative ad ora, giorno e mese della rilevazione climatica. Le prime, essendo caratterizzate da valori continui, sono state normalizzate mentre le seconde sono state codificate nelle loro componenti seno e coseno, in modo tale da preservarne la natura ciclica.

Per migliorare la qualità della previsione sono state aggiunte ulteriori informazioni ricavate dai questionari utilizzati nella fase di clustering. Queste riguardano:

1. La fascia oraria considerata "mattina", "pomeriggio", "sera" e "notte";
2. Il numero di occupanti presenti nell'abitazione per fascia oraria;
3. La tipologia del giorno considerato tra "weekend" e "weekday".

Gli attributi di tipo stringa sono stati codificati applicando la codifica one-hot encode mentre quelli di tipo numerico sono stati normalizzati prima di procedere con la fase di addestramento. Quando si fa previsione di serie storiche, due scelte che influenzano molto la qualità della previsione sono la finestra di addestramento e l'orizzonte di previsione. I dati storici dei consumi elettrici hanno come principale caratteristica quella di essere fortemente dipendenti da un punto di vista temporale. Prendendo in considerazione i consumi giornalieri di un'abitazione, è facile intuire che il loro andamento risulta stabile (nella media) se considerato in un intervallo di pochi minuti, mostrando delle variazioni significative considerando ore o fasce orarie differenti. La presenza di un picco di consumo, per quanto improvviso possa essere, risulta essere sempre preceduto da un innalzamento dei consumi negli istanti temporali precedenti. Quindi, ai fini della previsione, il modello può essere "aiutato" utilizzando per l'addestramento delle serie di consumi temporalmente consecutivi. Tali serie sono state chiamate "finestre di addestramento".

Data una serie storica di consumi $\{x_1, \dots, x_n\}$ e fissata la dimensione della finestra pari a $k > 0$, una generica finestra di addestramento è definita come $\{x_i, x_{i+1}, \dots, x_{i+k-1}\}$ e viene utilizzata per

prevedere il consumo x_{i+k} . La scelta della finestra per il dataset utilizzato è stata effettuata in crossvalidation, scegliendo nel range di valori da 1 (caso in cui la finestra non viene utilizzata) a 5.

La previsione dei consumi elettrici per essere significativa deve restituire i consumi predetti relativi ad un intervallo futuro chiamato "orizzonte di previsione". Fissato un orizzonte, la finestra dei consumi noti fino al tempo t viene utilizzata per predire il valore successivo x_{t+1} . Quest'ultimo viene aggiunto alla finestra in modo tale da essere utile per la previsione del valore successivo x_{t+2} , la finestra viene shiftata avanti di una posizione e si procede alla nuova previsione. Questa logica viene iterata per predire tutti i consumi nell'orizzonte fissato.

Data la serie storica di consumi $\{x_1, \dots, x_n\}$, fissato la dimensione dell'orizzonte di previsione pari ad $h > 0$, l'obiettivo è la previsione dei consumi futuri $\{x_{i+1}, \dots, x_{i+h}\}$.

Nell'analisi presentata sono stati ipotizzati due orizzonti di previsione per ogni granularità dei dati considerata, ovvero:

1. orizzonte pari a "sei" e "otto" per i consumi con campionatura di cinque minuti;
2. orizzonte pari a "quattro" e "sei" per i consumi con campionatura di quindici minuti.

I risultati ottenuti per i diversi casi sono stati utili per valutare l'impatto dell'orizzonte in termini di accuratezza della previsione. L'aspettativa è che l'accuratezza diminuisca all'aumentare dell'orizzonte utilizzato, dal momento che questo prevede l'utilizzo di consumi predetti come base per la predizione dei consumi successivi.

Il passo successivo alla costruzione del dataset è stata la scelta del modello da considerare. Il primo modello di addestramento è stato selezionato tra una serie di modelli caratterizzati da una complessità ridotta, visto il potenziale utilizzo in ambito residenziale dove non è realistico pensare di poter addestrare modelli complicati come reti neurali ricorrenti. Peraltro, diversi studi in letteratura, come in [39], hanno dimostrato che, nonostante la loro semplicità, è possibile ottenere delle buone previsioni utilizzando tali modelli. Si è deciso quindi di valutare in crossvalidation le performance dei modelli: SVM, Random Forest e Gradient Boosting. Per definire un orizzonte di previsione si è utilizzata la classe Regressor Chain [40] di Scikit Learn. Tale classe permette di creare una "catena" di modelli in cui il primo viene addestrato con le variabili indipendenti per predire il primo valore y_1 dell'orizzonte, il secondo verrà addestrato sulle variabili indipendenti e y_1 per predire y_2 , etc. Al termine della previsione di ogni orizzonte la finestra delle osservazioni viene fatta avanzare di una posizione per procedere alla previsione dell'orizzonte successivo. Tale approccio permette di ottenere diverse previsioni relative ad uno stesso istante temporale, si è deciso quindi di calcolarne il valore medio per definire la previsione restituita dal modello. La scelta dei modelli da utilizzare per la fase di previsione, il tuning dei loro parametri e la dimensione della finestra di addestramento sono state definite applicando la tecnica della cross validation, adattata per funzionare sulla previsione di serie temporali. Ipotizzando di avere un dataset di lunghezza n e di utilizzare finestre di dimensione k , il dataset può essere suddiviso in n/k porzioni. Per mantenere la correlazione temporale tra le osservazioni il training-set ed il validation-set sono stati selezionate porzioni consecutive. Nella prima iterazione si utilizzerà la prima porzione come training-set e la seconda come validation-set, nella seconda si utilizzeranno le prime due porzioni come training-set e la terza come validation-set e così via. L'idea è di incrementare la dimensione del training-set ad ogni iterazione, facendo slittare in avanti anche la selezione del validation-set. In questo modo è possibile variare entrambi i dataset senza far venire meno il fattore temporale che li caratterizza.

Per scegliere il modello da utilizzare è stato innanzitutto definito un benchmark iniziale sulla base delle previsioni ottenute da un modello di bassa complessità. Successivamente, per migliorare i

risultati ottenuti, è stata implementata una rete neurale. Dalla letteratura [39] è emerso che i modelli più semplici con cui si sono ottenuti i migliori risultati negli studi degli ultimi anni, relativi alla previsione dei consumi elettrici, sono la Support Vector Machine, il Random Forest e il Gradient Boosting.

Nello specifico quest'ultimo è il modello che generalmente restituisce i risultati migliori. I modelli sono stati valutati in cross-validation con una finestra variabile tra 0 e 5 utilizzando i dati relativi all'abitazione C7. Tale scelta è dovuta a due caratteristiche dell'abitazione:

- è il "centroide" del suo cluster, cioè l'elemento più rappresentativo;
- è caratterizzata da consumi medi e quindi da un profilo energetico nella media.

Per valutare i risultati ottenuti, si devono introdurre delle metriche adatte a valutare la qualità della previsione. In particolare si sono adottate le seguenti metriche:

Dato un valore reale y_i e la relativa previsione \hat{y}_i , l'errore e_i è definito come la differenza tra valore vero e valore predetto:

$$e_i = y_i - \hat{y}_i$$

Per valutare la bontà di un modello spesso non è possibile affidarsi ad un solo indicatore a causa di diversi aspetti da tenere in considerazione.

Nell'analisi presentata le previsioni sono state valutate rispetto all'andamento medio dei consumi ed i picchi di consumo. Le metriche utilizzate sono:

1. *Root Mean Square Error* (RMSE): è definito come l'errore quadratico medio:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

La presenza del quadrato lo rende sensibile alla presenza di outliers, motivo per cui è stato utilizzato per valutare la bontà delle previsioni rispetto ai picchi di consumo.

2. *Mean Absolute Error* (MAE): è definito come l'errore medio sul campione. Ha la caratteristica di trattare equamente tutti gli errori, sia quelli più significativi che meno. A differenza del precedente non è sensibile agli *outlier* ed è utile per determinare la bontà delle previsioni rispetto all'andamento medio dei consumi.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

3. *Mean Absolute Percentage Error* (MAPE): a differenza dei due indicatori precedenti, quest'ultimo non è sensibile alla scala. Si dimostra quindi utile per confrontare i risultati ottenuti per dataset diversi. Vista la presenza di valori dei consumi nulli al denominatore, si è preferito l'utilizzo dell'indicatore *Symmetric MAPE* (SMAPE). Quest'ultimo può assumere valori tra 0% e 200%.

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

$$SMAPE = \frac{200\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)}$$

Nella valutazione della bontà delle previsioni si è scelto di dare un peso maggiore alla metrica RMSE con l'obiettivo di ottenere l'accuratezza migliore possibile rispetto alla previsione dei picchi, aspetto considerato più importante rispetto all'andamento medio dei consumi.

Le Random Forest sono un metodo di apprendimento di tipo "ensemble" per la classificazione e la regressione che operano costruendo, in modo indipendente, una moltitudine di alberi decisionali durante l'addestramento. I risultati ottenuti dagli alberi di decisione vengono poi combinati con la media (nel caso di problemi di regressione) o utilizzando il voto a maggioranza (nel caso di problemi di classificazione).

Il Gradient Boosting [42] è un altro metodo di ensemble che combina l'utilizzo di più modelli al fine di migliorarne l'efficacia individuale. Nello specifico, l'ensemble learning è un paradigma di addestramento basato sull'utilizzo di più modelli semplici (weak learners), con l'obiettivo di combinarli opportunamente ed ottenere un modello più espressivo (strong learner). Tra i metodi di aggregazione più utilizzati c'è il boosting, che è una tecnica di ensemble learning basata sulla costruzione di diversi alberi di decisione [43]. Il principio alla base del boosting è quello di potenziare un modello semplice (cioè modelli che sul training set hanno errore elevato) creando nuovi modelli che vengono addestrati per correggere gli errori effettuati dai modelli precedenti sul training set. A differenza delle Random Forest, durante l'addestramento, gli alberi non sono indipendenti ma utilizzano le informazioni relative agli addestramenti precedenti. Così come gli altri metodi di boosting, il modello viene costruito in modo graduale, ma il Gradient Boosting generalizza gli altri metodi consentendo l'ottimizzazione di una funzione di costo differenziabile arbitraria tramite il metodo di discesa del gradiente.

Come modello più complesso è stata utilizzata una rete LSTM che è una rete neurale ricorrente capace di catturare le dipendenze temporali tra i dati in ingresso. La rete Long-Short Term Memory (LSTM) [44] è una variante della rete ricorrente introdotta da Sepp Hochreiter e Juergen Schmidhuber nel 1997 con la caratteristica di considerare le dipendenze temporali di lungo termine. Da un punto di vista architetturale la differenza tra le due reti riguarda la gestione delle informazioni in memoria. Il neurone della LSTM, rappresentato nella Figura 15 ha la possibilità di aggiungere o eliminare informazioni dalla propria cella di memoria. Per rendere possibili queste operazioni è dotato di una struttura più complessa, caratterizzata dalla presenza di tre elementi, detti gate, utili alla gestione dello stato chiamati input gate, output gate e forget gate.

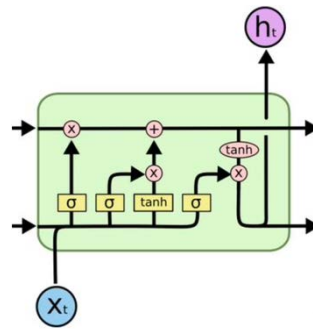
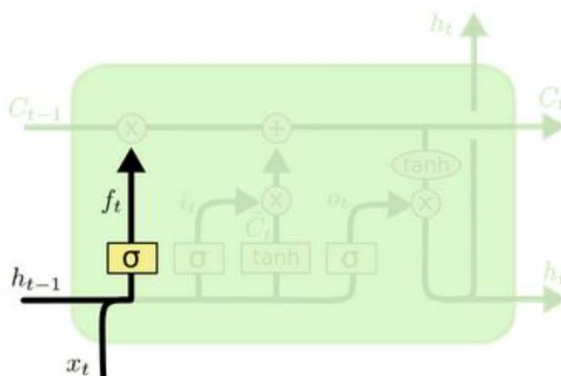


Figura 15: Neurone della rete LSTM

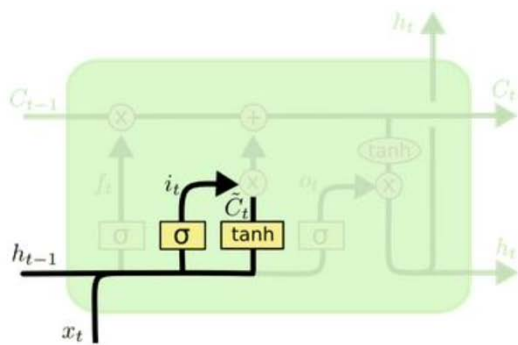
Il forget gate, in Figura 16, decide le informazioni che verranno rimosse dallo stato del neurone. Riceve in input l'informazione corrente (x_t) e l'output del neurone precedente (h_{t-1}), quindi decide la porzione dell'informazione da mantenere in memoria. Questo è possibile attraverso l'utilizzo della funzione sigmoidale $\sigma(h) = \frac{1}{1+e^{-h}}$ che, per definizione, genera valori nell'intervallo $[0,1]$. Moltiplicare un'informazione per zero ed uno corrisponde rispettivamente ad eliminarla e mantenerla in memoria.



$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Figura 16: Forget gate

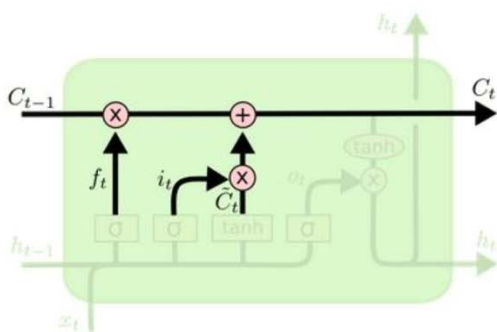
L'input gate, in Figura 17-18, ha il compito di aggiornare lo stato del neurone in base alle nuove informazioni arrivate in ingresso. Questo avviene in due fasi. Prima vengono definiti i valori dello stato che verranno aggiornati e quali informazioni in ingresso verranno salvate (i_t). Vengono generati tutti i possibili valori da aggiungere allo stato ($C'(t)$) e si procede ad aggiornare il vecchio stato attraverso le nuove informazioni.



$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_f)$$

$$f_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

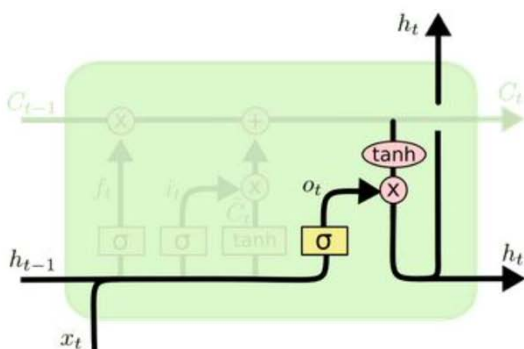
Figura 17: Input gate fase 1



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_{t-1}$$

Figura 18: Input gate fase 2

L'output gate, rappresentato in Figura 19, ha il compito di definire la porzione dello stato che verrà inviata in uscita, la quale è chiamata hidden state (h_t) .



$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$f_t = o_t * \tanh(C_t)$$

Figura 19: Output gate

L'architettura finale della rete utilizzata per predire i consumi elettrici è la seguente ed è descritta in Figura 20:

1. Input layer: un layer LSTM con units=100 e dropout=0.1;
2. Hidden layer: un layer Dense con 25 unità e funzione di attivazione ReLu;
3. Output layer: realizzato inserendo un ulteriore layer Dense con un numero di unità pari all'orizzonte di previsione e funzione di attivazione ReLu.



Figura 20: Architettura della rete LSTM utilizzata

L'analisi è stata implementata utilizzando il linguaggio Python. Tale scelta è legata alla disponibilità di numerose librerie open-source ampiamente utilizzate per analisi nell'ambito data-science e machine-learning. In particolare, sono state utilizzate le implementazioni di k-means, random forest e gradient boosting presenti all'interno del package scikit-learn. La rete neurale è stata implementata utilizzando le API offerte dalla libreria Keras.

3.4 Risultati della previsione

Ogni modello è stato addestrato utilizzando i dati di ogni abitazione identificata come centroide del proprio cluster di appartenenza. Il modello addestrato è stato quindi utilizzato per predire i consumi di tutte le abitazioni appartenenti al medesimo cluster. Per ogni abitazione l'80% del dataset a disposizione è stato utilizzato come training-set, il restante 20% come test-set. Le previsioni ottenute per i consumi con campionamento di 5 minuti sono calcolate per orizzonte di previsione 6 e 8. Per il campionamento a 15 minuti l'orizzonte di previsione è 4 e 6.

Nella fase di predizione si sono osservati i diversi comportamenti dei due modelli utilizzati, cioè Gradient Boosting e rete neurale LSTM. In particolare, la rete neurale ha mostrato di saper prevedere in maniera efficiente l'andamento dei consumi, mostrando però dei limiti di accuratezza nella previsione del valore dei picchi. Al contrario il Gradient Boosting ha come principale punto di forza la maggior accuratezza nella predizione dei picchi, risultando però meno accurato riguardo l'andamento dei consumi. Come naturale conseguenza si è cercato di "unire" le predizioni dei due modelli con l'obiettivo di sfruttare i punti di forza di entrambi.

L'idea è stata implementata attraverso le seguenti fasi:

- Definizione di una soglia attraverso cui identificare i picchi di consumo
- Implementazione di un classificatore e predizione dei picchi
- Creazione di una predizione di tipo "ensemble"

La classificazione dei picchi è basata sulle previsioni ottenute dal Gradient Boosting in quanto più accurate rispetto ai picchi. Tali previsioni hanno però la caratteristica di sottostimare in generale il valore dei picchi. Per migliorare la classificazione la previsione è stata moltiplicata per un peso, il "fattore di scala", con l'obiettivo di far avvicinare il valore dei picchi predetti a quello dei picchi reali. Tale valore è stato determinato in cross-validation partendo da un peso iniziale pari a 1.1 e

incrementandolo ad ogni iterazione di 0.1. L'individuazione dei picchi è stata effettuata analizzando i consumi su intervalli di ventiquattro ore. Calcolando un "valore di soglia" per ogni intervallo, i consumi sono stati suddivisi in due classi:

- superiori alla soglia, cioè i picchi;
- inferiori alla soglia, ovvero il resto dei consumi.

3.4.1 Previsioni a 5 minuti

La classificazione effettuata per i centroidi ha restituito i seguenti risultati. Si può osservare che l'accuratezza nella predizione dei picchi di consumo cresce sensibilmente passando dalle abitazioni C3-C6 a C4-C2. Tale effetto può essere ricondotto alla natura dei cluster. I primi due sono infatti caratterizzati da consumi più eterogenei e la previsione dei picchi, di conseguenza, risulta essere più difficile.

	F1 score – classe 0	F1 score – classe 1	Accuratezza	Macro Avg	Weighted Avg
C3	0.92	0.89	0.91	0.91	0.91
C6	0.93	0.88	0.91	0.91	0.91
C4	0.93	0.93	0.93	0.93	0.93
C2	0.96	0.96	0.96	0.96	0.96

Tabella 12- Risultati ottenuti nella classificazione dei picchi di consumo con dati aggregati a 5 minuti

1. Per il cluster 1 si hanno i seguenti risultati:

	RMSE	MAE	SMAPE
C3	201.665519	114.6127305	61.48943565

Tabella 13- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 1 con orizzonte 6

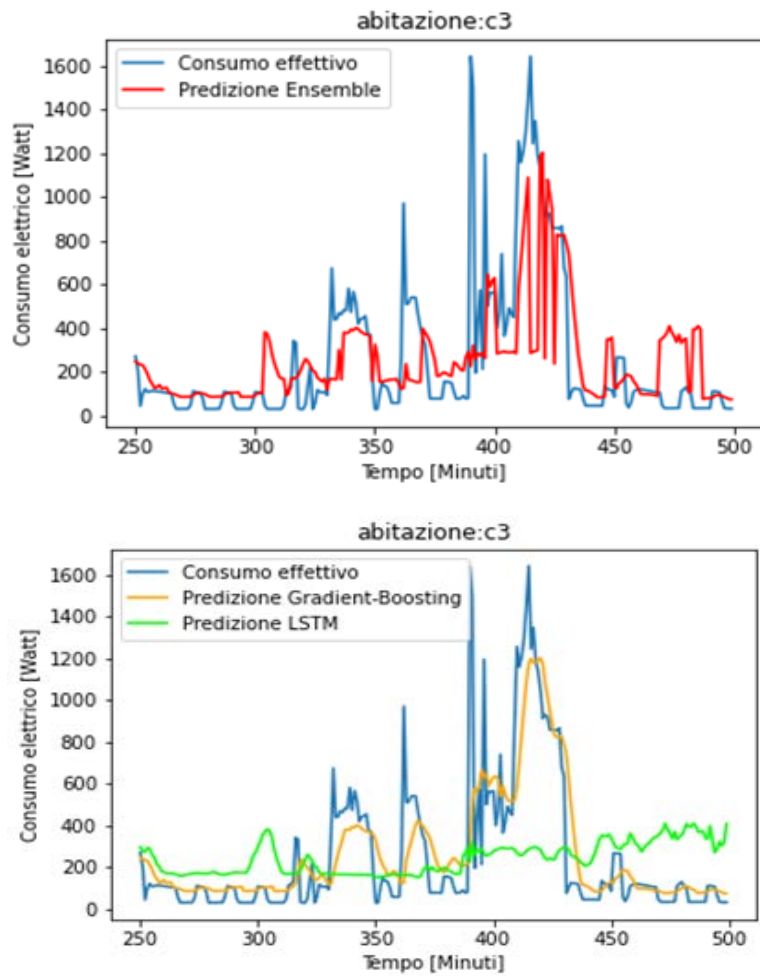


Figura 21: plot della previsione sul test set per la casa C3

2. Per il cluster 2 si hanno i seguenti risultati:

	RMSE	MAE	SMAPE
C6	255.542753	109.314807	49.990913
C1	195.349693	95.103113	27.370190
C5	273.237997	121.074462	32.515245
C7	316.339834	125.040883	51.729216

Tabella 14- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 2 con orizzonte 6

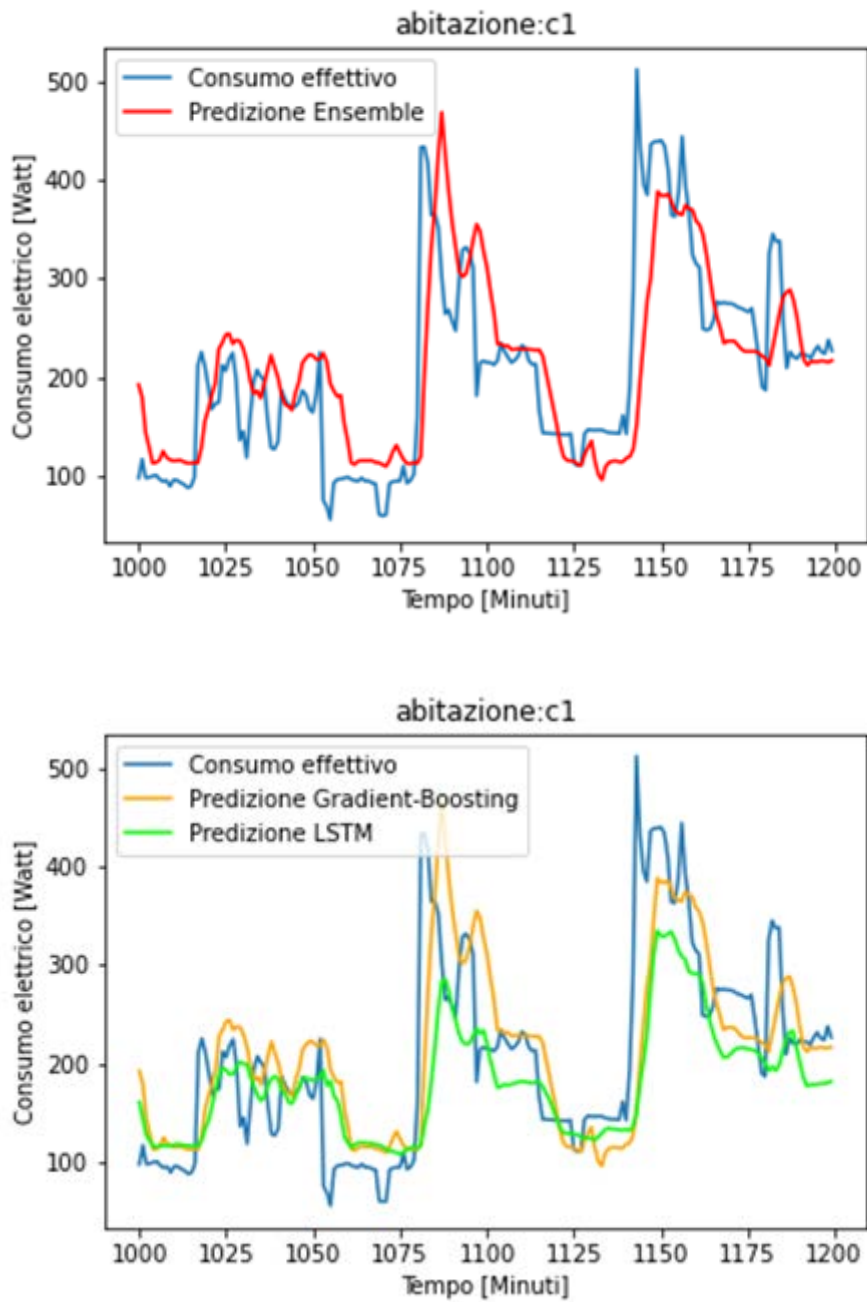


Figura 22: plot della previsione sul test set per abitazione C1

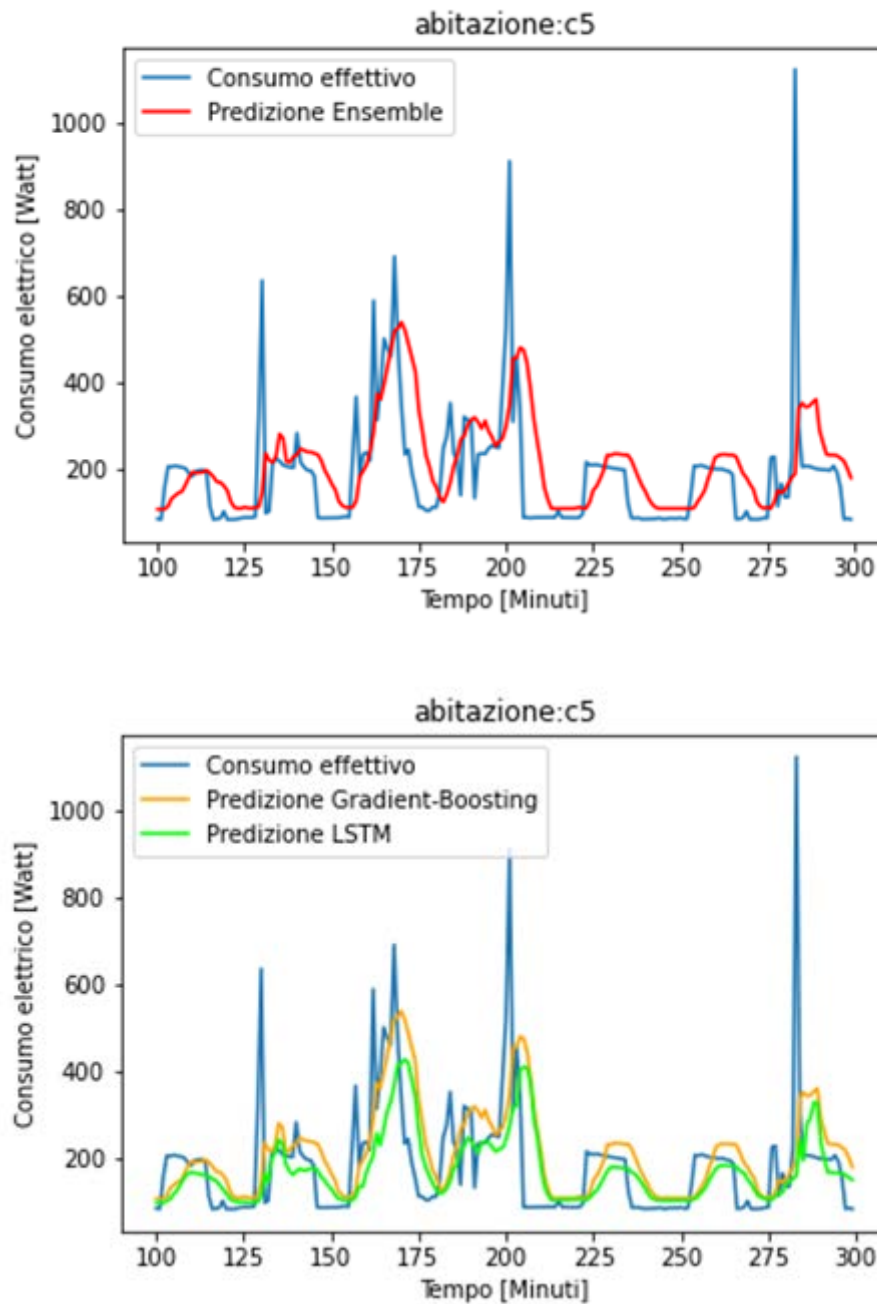


Figura 23: plot della previsione sul test set per abitazione C5

plot della previsione sul test set per la casa C3

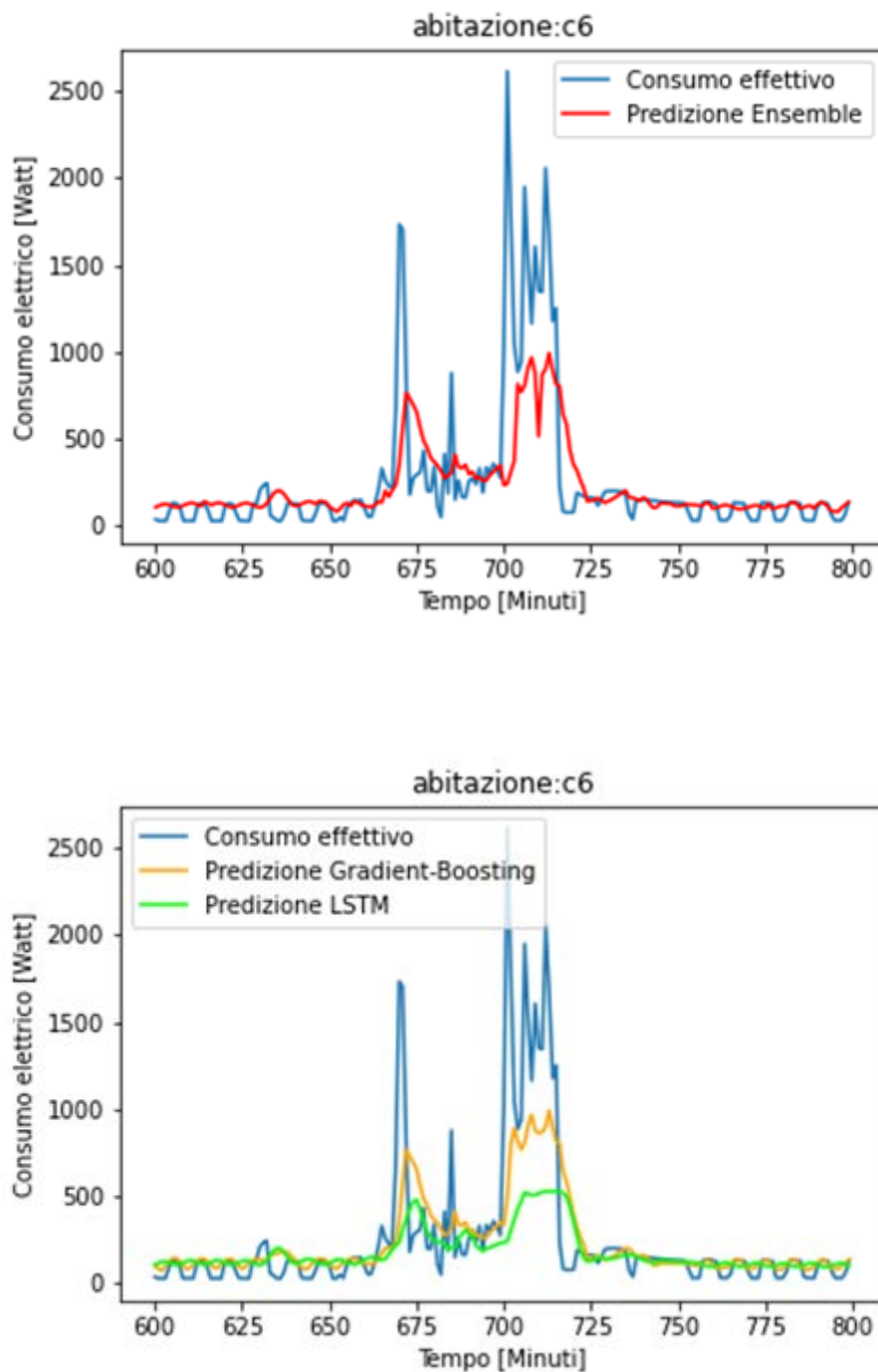


Figura 24: plot della previsione sul test set per abitazione C6

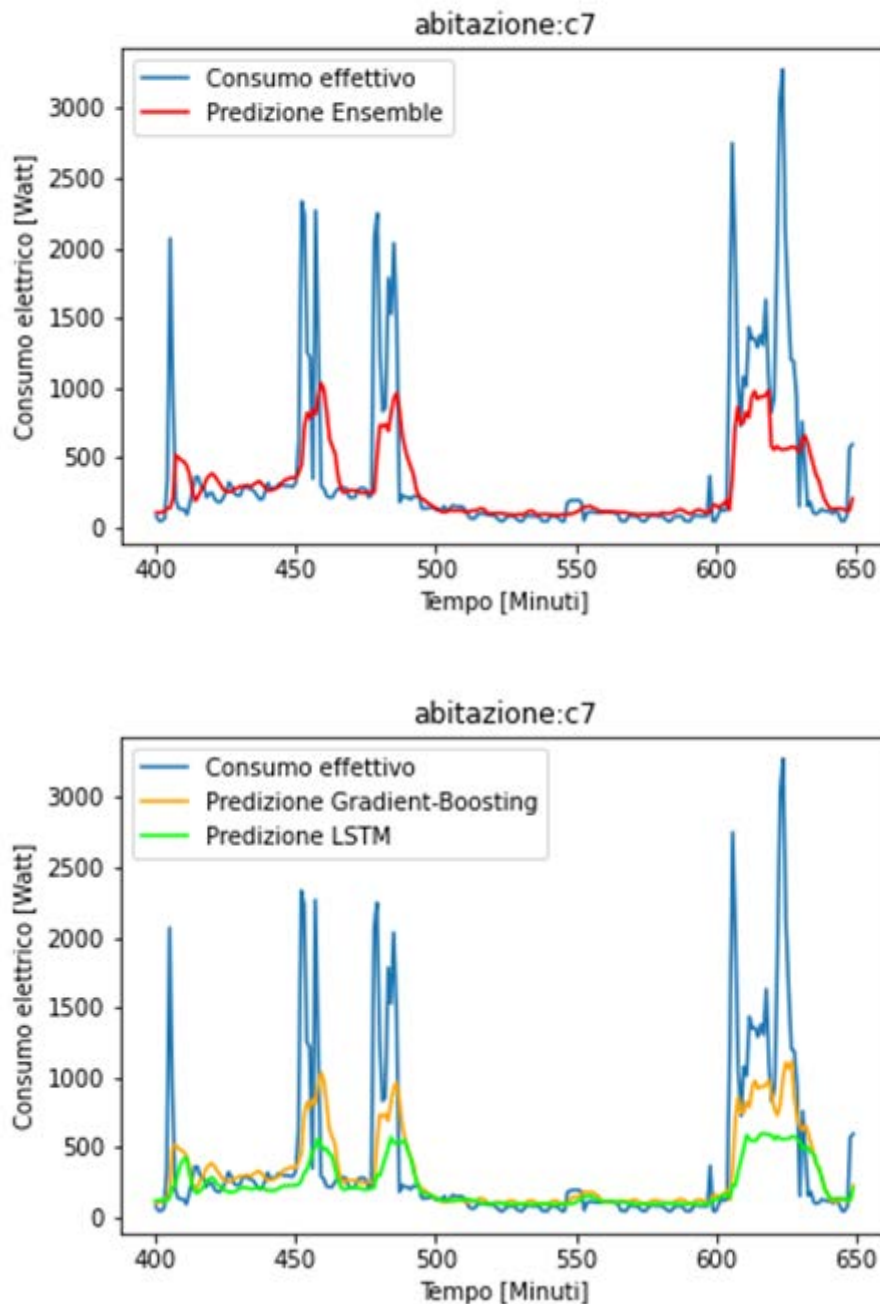


Figura 25: plot della previsione sul test set per abitazione C7

3. Per il cluster 3 si hanno i seguenti risultati:

	RMSE	MAE	SMAPE
C4	189.48376	82.0272467	61.49221243

Tabella 15- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 3 con orizzonte 6

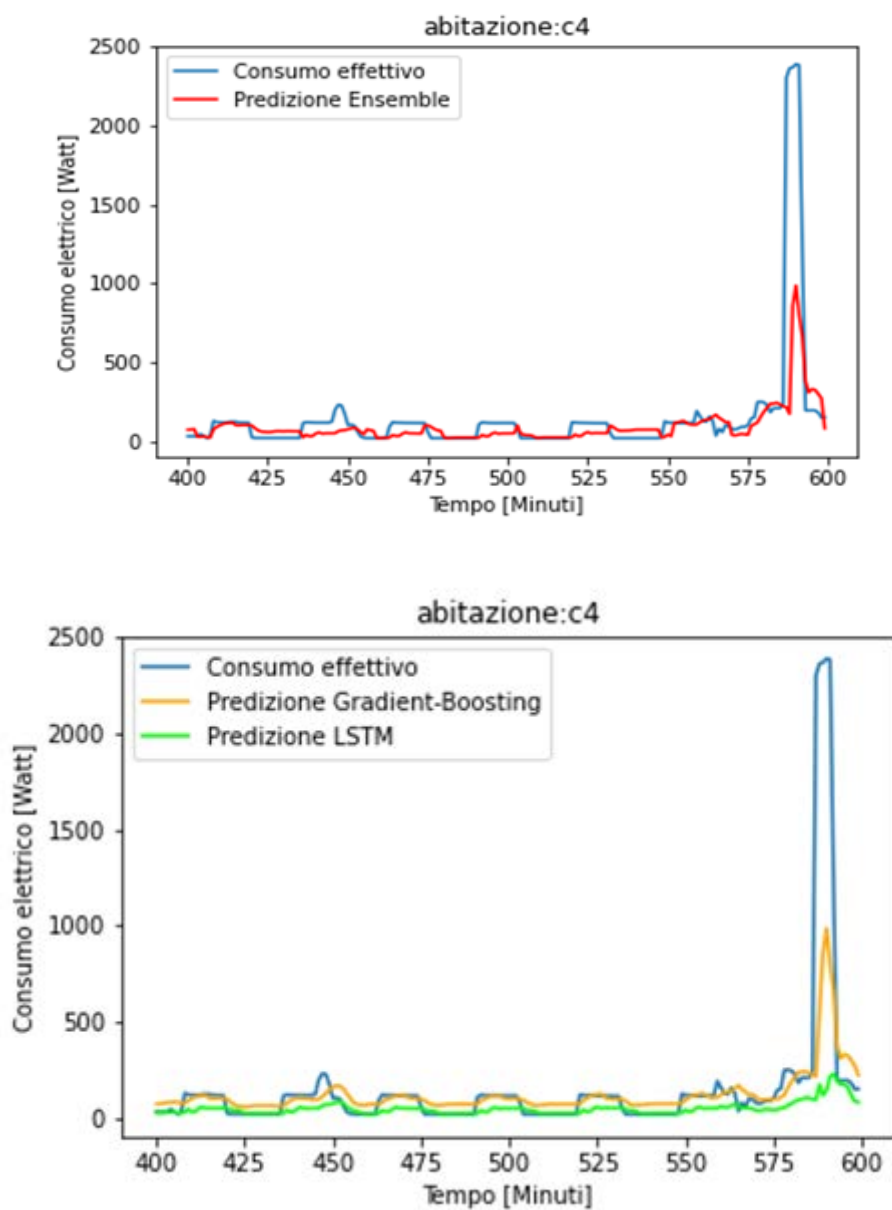


Figura 26: plot della previsione sul test set per abitazione C4

4. Per il cluster 4 si hanno i seguenti risultati:

	RMSE	MAE	SMAPE
C2	224.437761	79.86417272	47.92640563
C8	168.426021	76.11657569	66.66831718

Tabella 16- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 4 con orizzonte 6

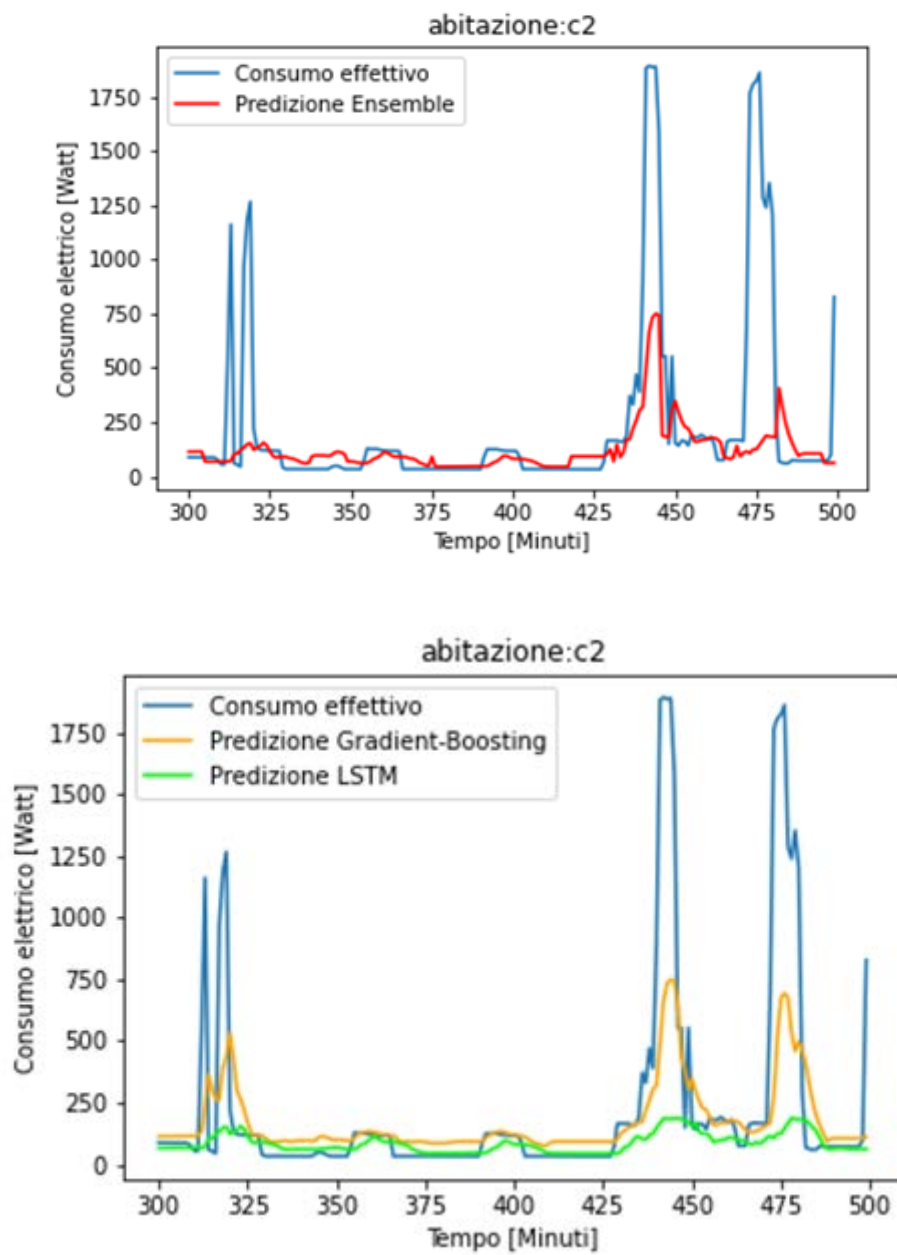


Figura 27: plot della previsione sul test set per abitazione C2

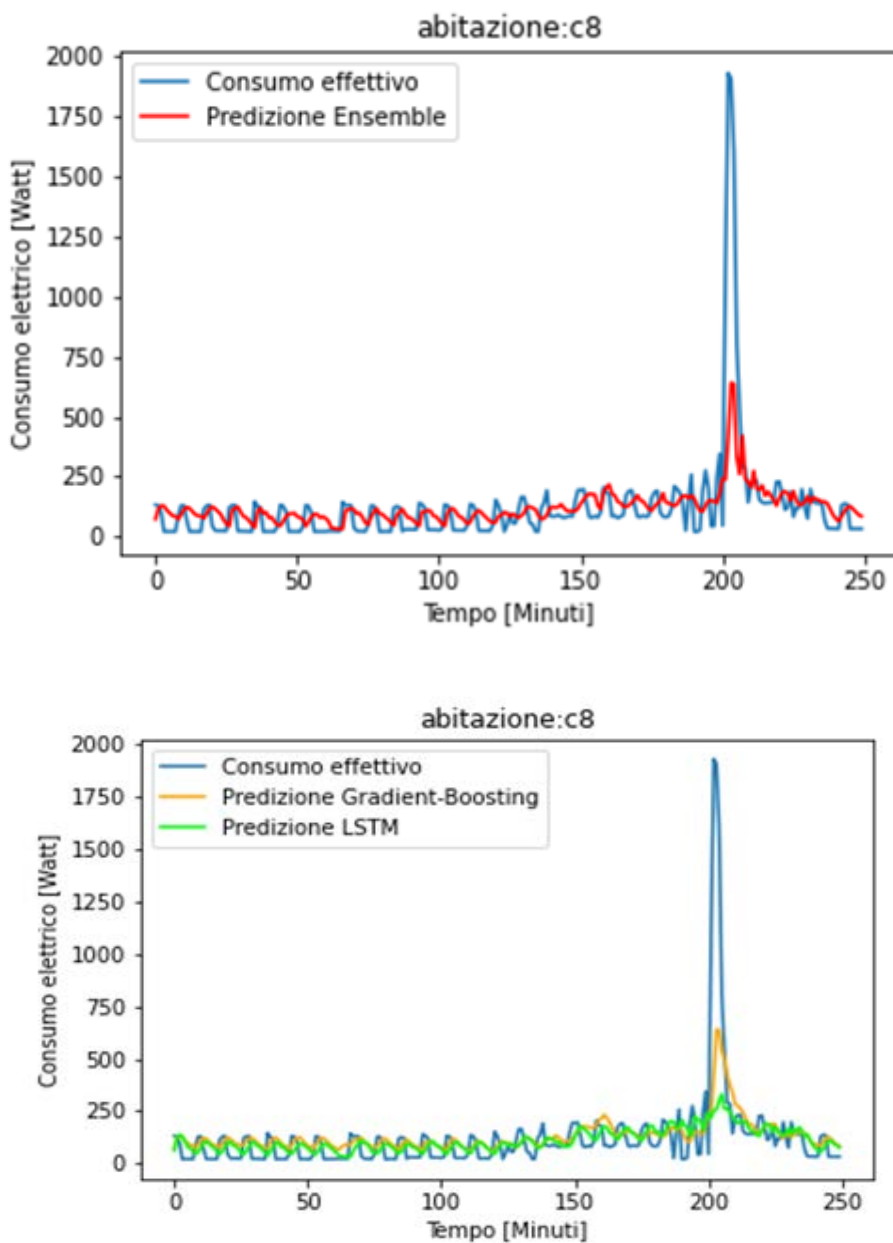


Figura 28: plot della previsione sul test set per abitazione C8

Se invece si sceglie orizzonte di previsione pare a 8 si ottengono i seguenti risultati:

1. Per il cluster 1 :

	RMSE	MAE	SMAPE
C3	205.8566314	120.2594428	64.28809927

Tabella 17- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 1 con orizzonte 8

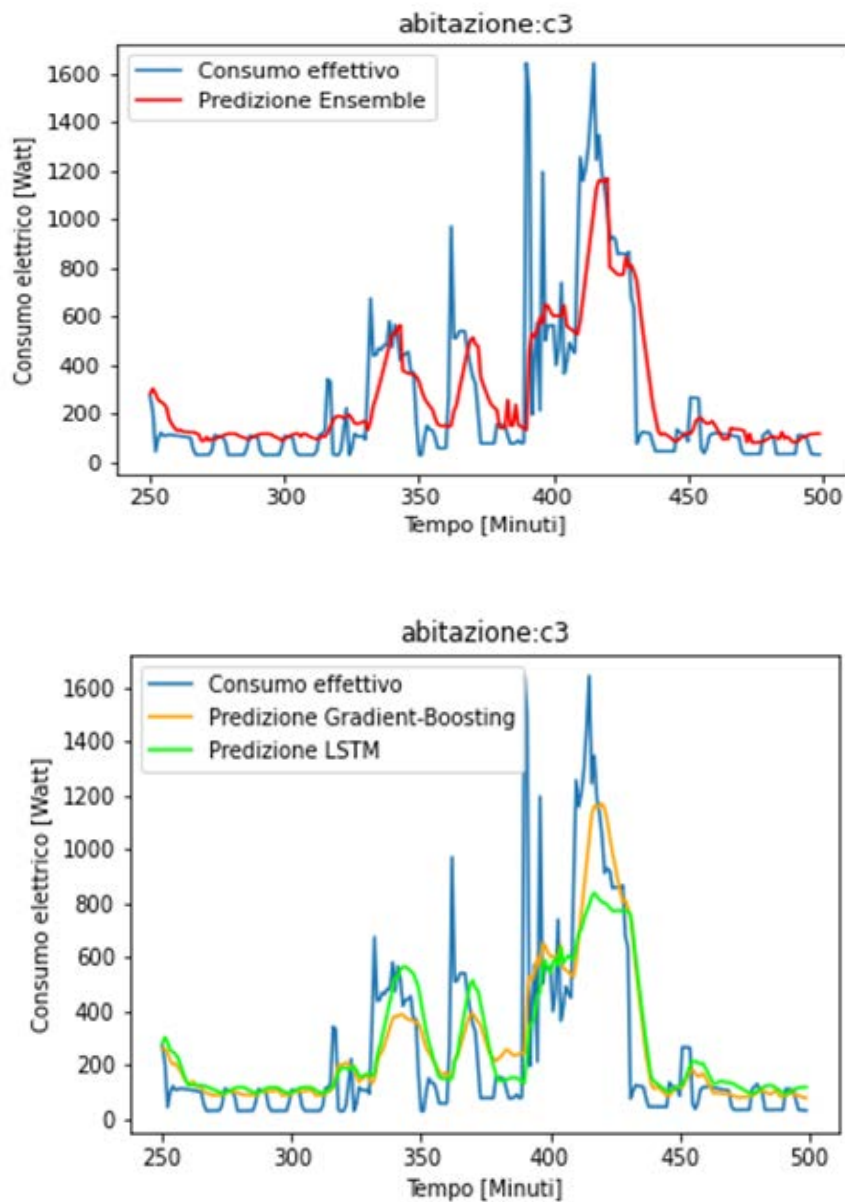


Figura 29: plot della previsione sul test set per abitazione C3

2. Per il cluster 2:

	RMSE	MAE	SMAPE
C6	260.772738	113.081444	50.607024
C1	205.424654	102.842986	29.452388
C5	284.686090	131.440133	36.133458

C7	331.551347	131.913938	53.397354
----	------------	------------	-----------

Tabella 18- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 2 con orizzonte 8

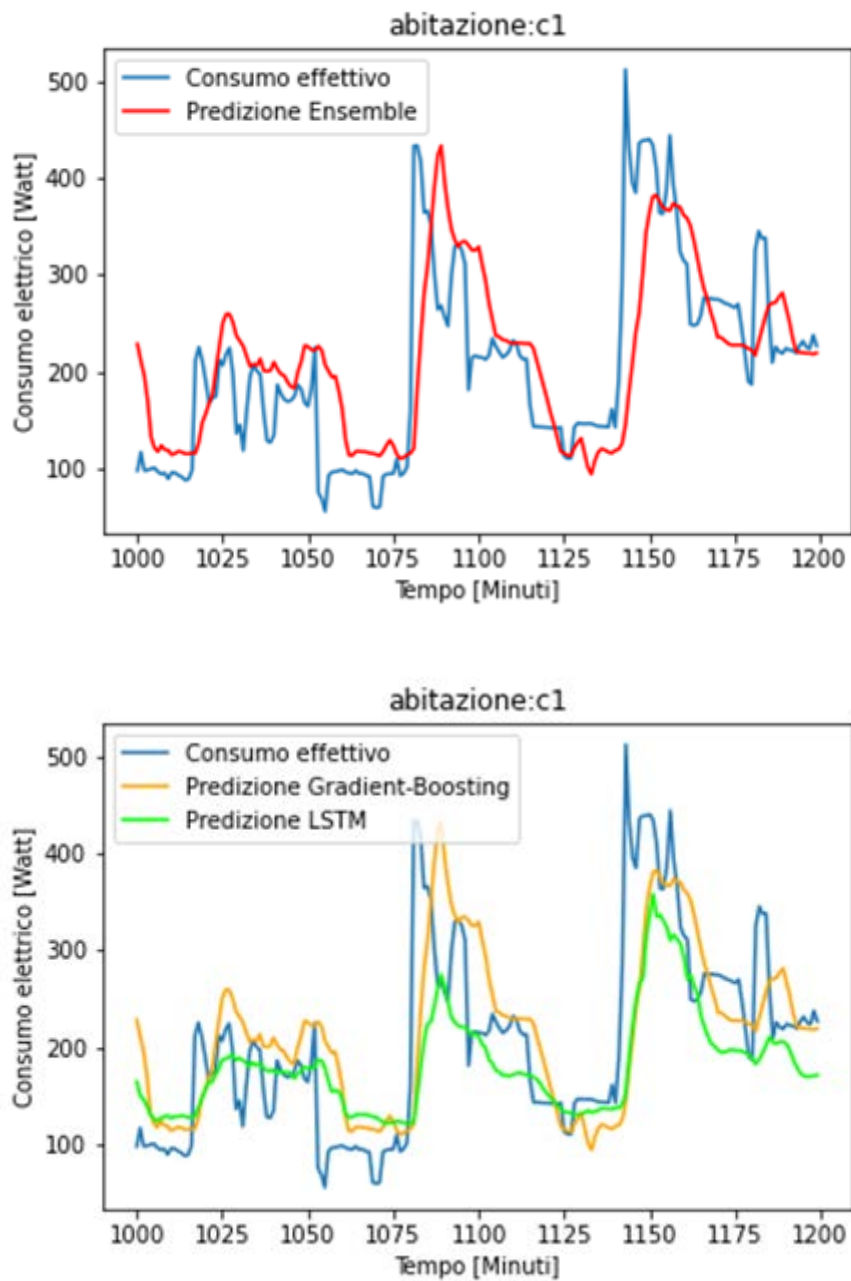


Figura 30: plot della previsione sul test set per abitazione C1

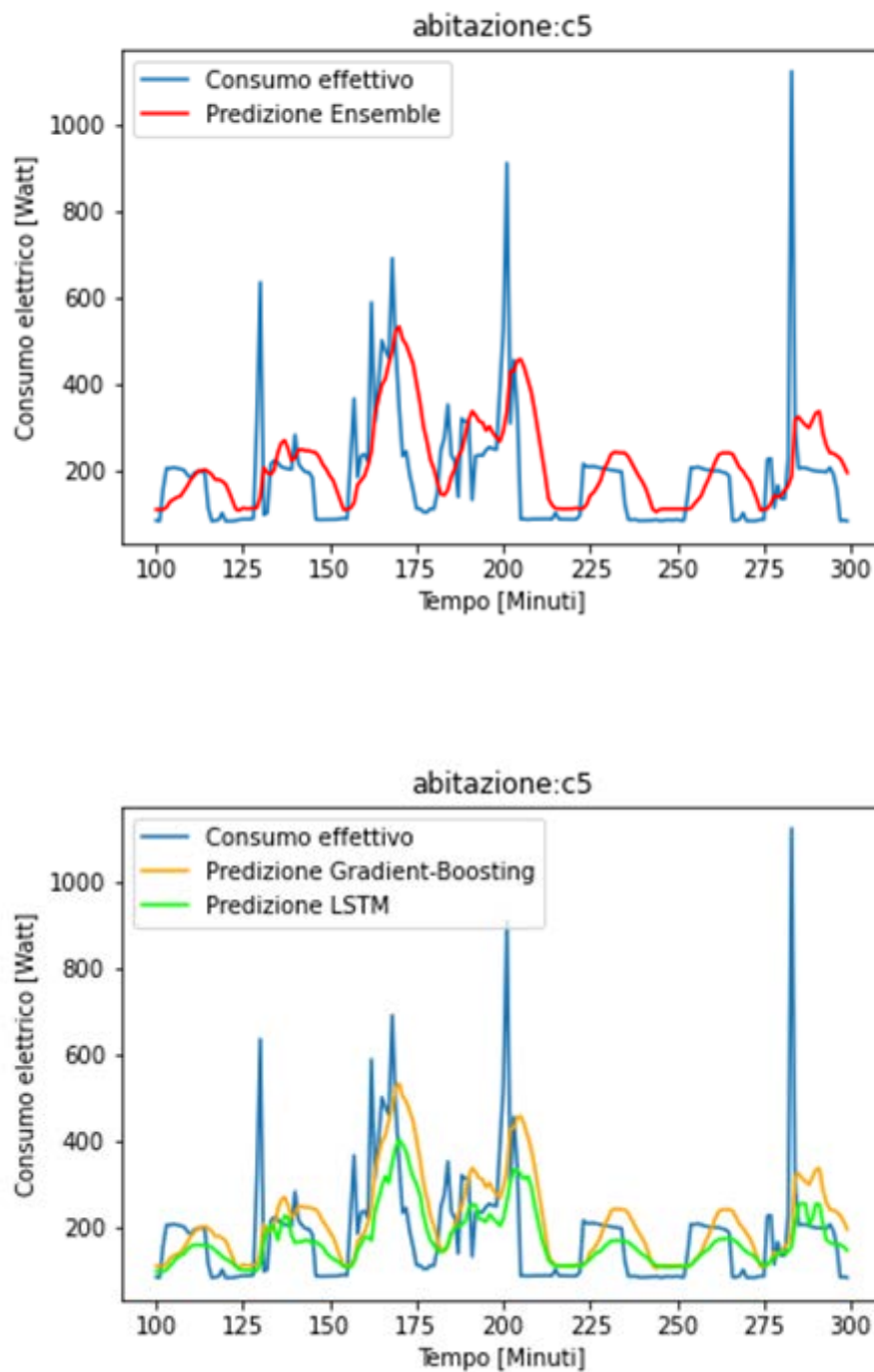


Figura 31: plot della previsione sul test set per abitazione C5

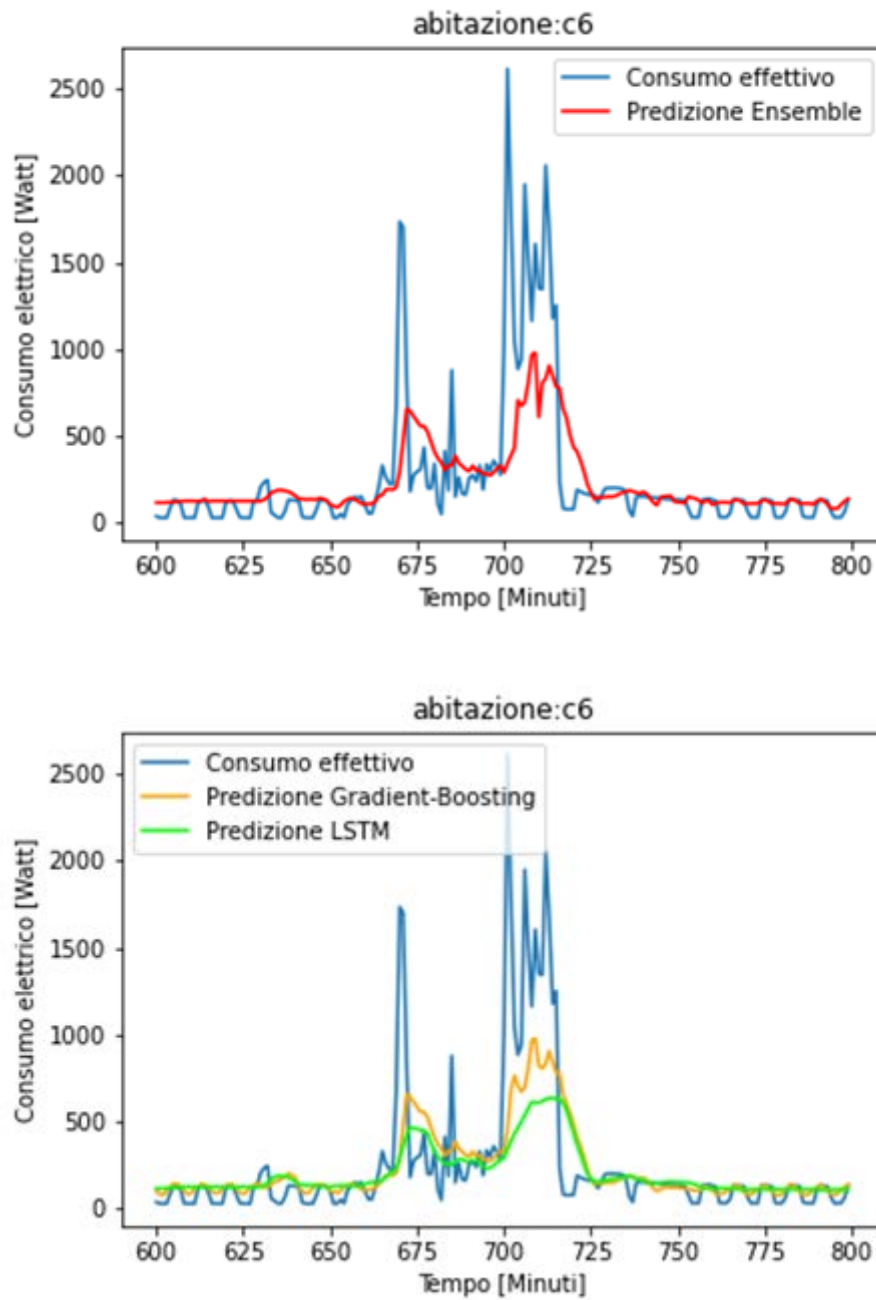


Figura 32: plot della previsione sul test set per abitazione C6

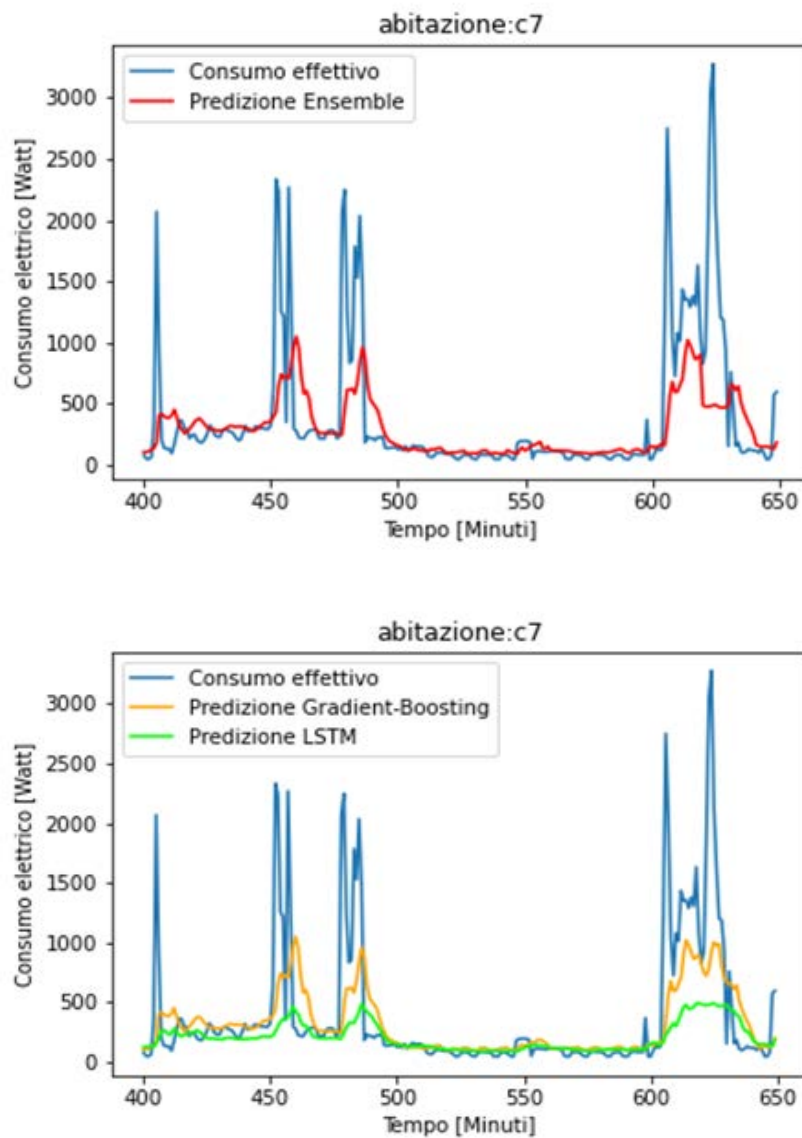


Figura 33: plot della previsione sul test set per abitazione C7

3. Per il cluster 3 si hanno i seguenti risultati:

	RMSE	MAE	SMAPE
C4	186.214764	77.094574	56.468062

Tabella 19- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 3 con orizzonte 8

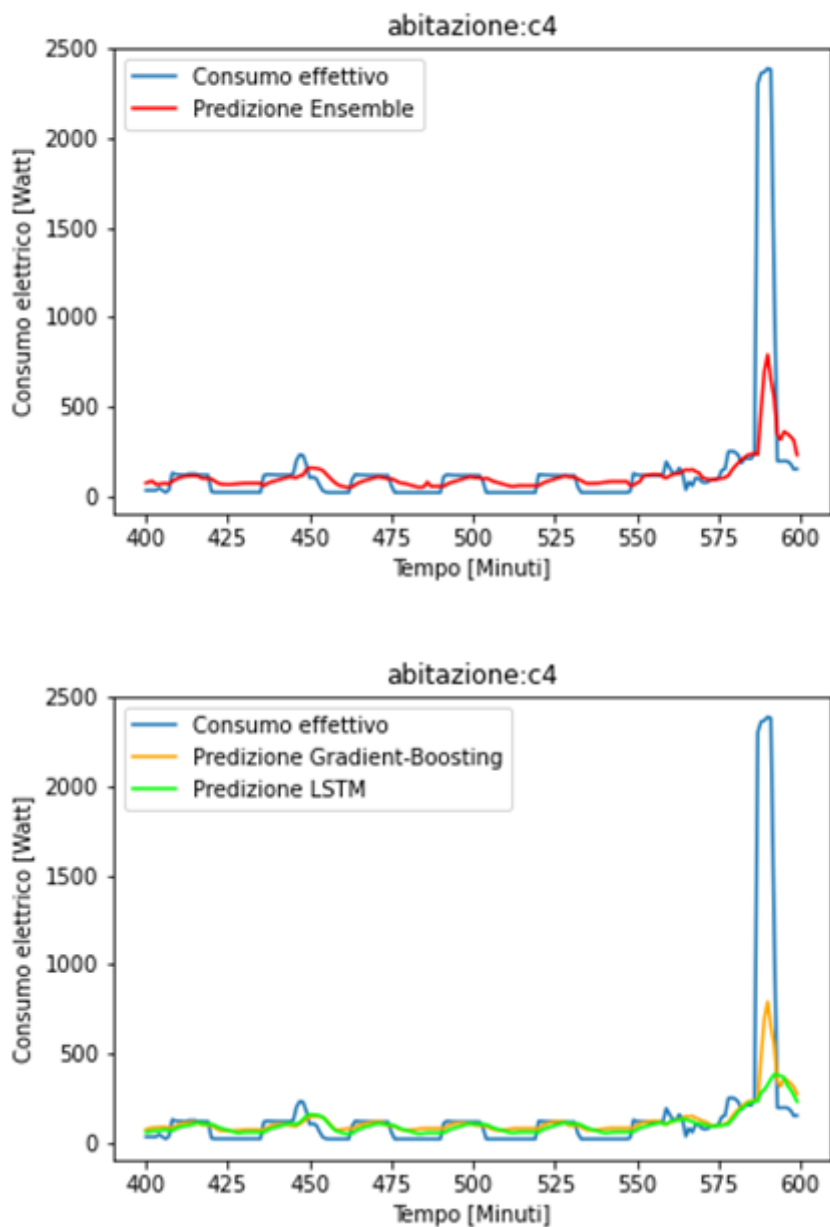


Figura 34: plot della previsione sul test set per abitazione C4

4. Per il cluster 4 si hanno i seguenti risultati:

	RMSE	MAE	SMAPE
C2	219.441100	83.016484	51.885850
C8	172.303903	73.824829	66.271234

Tabella 20- Risultati ottenuti nella previsione con dati aggregati a 5 minuti per il cluster 4 con orizzonte 8

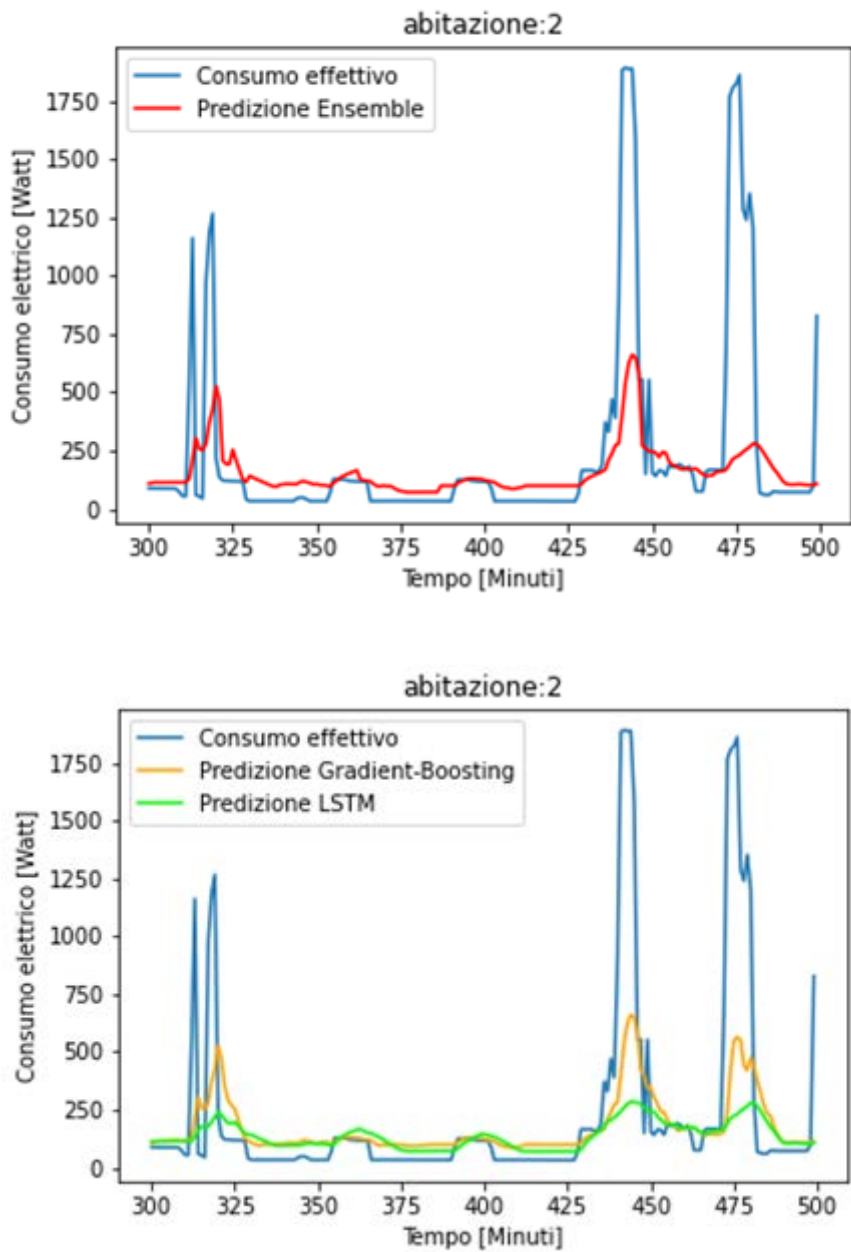


Figura 35: plot della previsione sul test set per abitazione C2

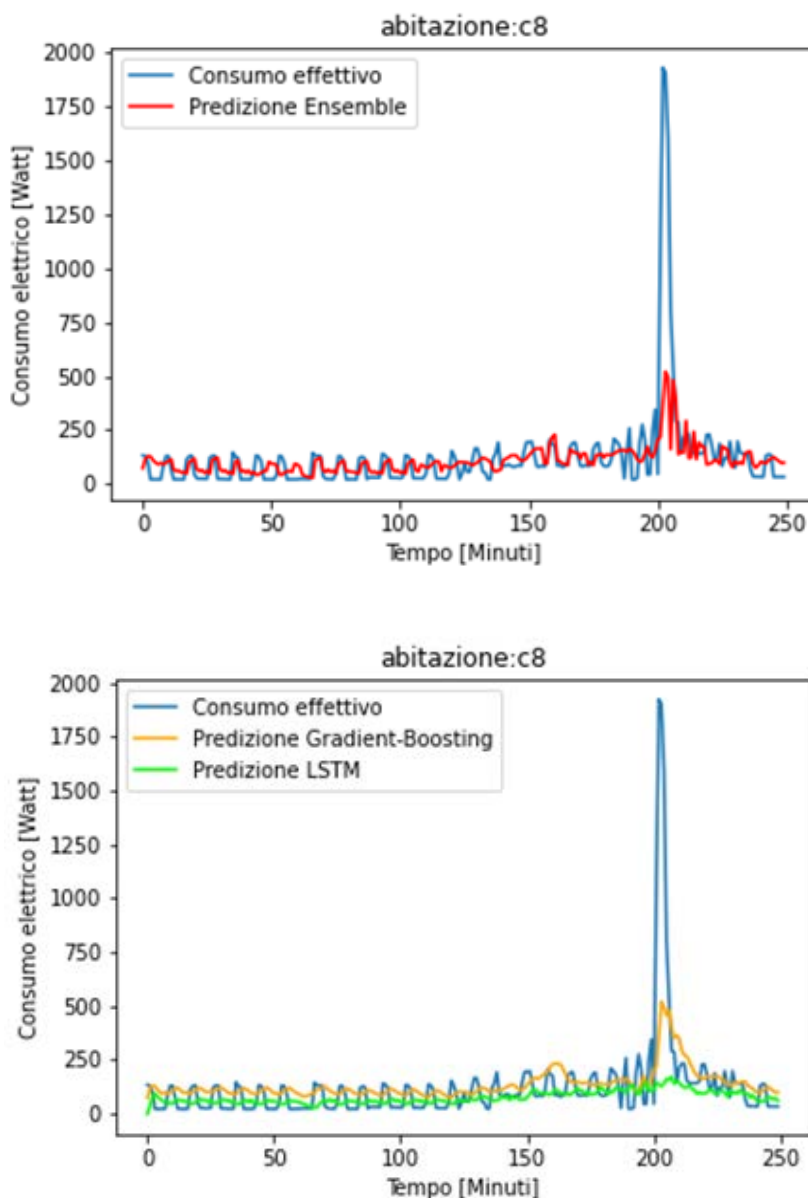


Figura 36: plot della previsione sul test set per abitazione C8

3.4.2 Confronto tra previsioni LSTM ed Ensemble Learning

I risultati ottenuti dall'Ensemble Learning sono stati confrontati con quelli ottenuti dalla rete neurale. Tale scelta è sembrata più corretta in quanto le previsioni del gradient boosting sono caratterizzate da un mae significativamente minore rispetto alla rete neurale. Spesso, tali previsioni si discostano parecchio dall'andamento reale dei consumi, risultando poco utilizzabili in un caso reale. L'ensemble learning è stato quindi "interpretato" come un miglioramento della previsione della rete neurale.

Nella tabella seguente sono state evidenziati i casi in cui l'ensemble learning permette di migliorare la previsione della rete neurale. Per i centroidi, nel caso della rete neurale, sono stati riportati i risultati relativi alle run migliori.

In generale si può osservare che l'ensemble learning riesce a migliorare la metrica rmse. In alcuni casi, come per l'abitazione C3, tale risultato è accompagnato da un peggioramento del mae e quindi dello smape. Questo effetto è conseguenza del classificatore introdotto per i picchi di consumo. Nei casi in cui tale predizione risulta essere sbagliata viene utilizzata, in assenza di un reale picco, la predizione del gradient boosting, introducendo una predizione dei consumi meno accurata rispetto alla rete neurale. Tale effetto risulta avere un impatto poco significativo in termini di mae ma è necessario visto il buon impatto in termini di rmse.

Abitazioni	Metrica	<u>Orizzonte 6</u>	<u>Orizzonte 6</u>	<u>Orizzonte 8</u>	<u>Orizzonte 8</u>
		<u>Ensemble</u>	<u>LSTM</u>	<u>Ensemble</u>	<u>LSTM</u>
C3	RMSE	201.6655191	204.3301685	205.8566314	210.3866303
	MAE	114.6127305	107.9644973	120.2594428	120.3192442
	SMAPE	61.48943565	32.28742398	64.28809927	65.07654074
C6	RMSE	255.5427526	294.5079641	260.7727384	282.4309927
	MAE	109.3148066	120.543095	113.0814441	120.1209602
	SMAPE	49.99091335	55.48102589	50.60702399	54.04159935
C1	RMSE	195.3496932	235.6149024	205.4246539	250.1204605
	MAE	95.10311347	106.2518499	102.8429862	112.9830696
	SMAPE	27.37019047	28.05602946	29.45238828	29.54421754
	RMSE	273.2379971	314.4173692	284.6860903	325.5991094

C5	MAE	121.0744625	123.9765915	131.4401326	128.574327
	SMAPE	32.51524481	32.34688	36.13345758	34.58356524
C9	RMSE	316.3398343	349.3553389	331.5513465	359.8253938
	MAE	125.040883	121.9657839	131.9139377	123.8639135
C7	SMAPE	51.72921597	46.65294208	53.39735364	47.95304295
	RMSE	189.4837598	197.1303777	186.2147645	196.9644523
C4	MAE	82.0272467	76.50811675	77.09457433	79.68348745
	SMAPE	61.49221243	49.93606252	56.4680619	55.55131957
C2	RMSE	224.437761	230.3545609	219.4411002	237.9767661
	MAE	79.86417272	67.2508304	83.01648449	82.31887778
C8	SMAPE	47.92640563	34.23834145	51.88584983	49.79612034
	RMSE	168.4260213	178.7319763	172.3039025	188.9876007
C8	MAE	76.11657569	75.23770587	73.82482871	72.96076379
	SMAPE	66.66831718	67.17520209	66.2712341	67.26126862

Tabella 21- Confronto tra ensemble learning e LSTM

3.4.3 Previsioni a 15 minuti

La classificazione e l'ensemble learning sono stati riprodotti in modo analogo per i consumi aggregati a 15 minuti. In generale i risultati ottenuti hanno un'accuratezza minore rispetto al caso precedente. Rimane invariato invece il rapporto tra i risultati dei primi due cluster rispetto agli ultimi due. Questi ultimi, essendo caratterizzati da consumi più omogenei, permettono di classificare i picchi con una qualità maggiore. I risultati ottenuti per la classificazione sono riportati nella seguente tabella.

	F1 score – classe 0	F1 score – classe 1	Accuratezza	Macro Avg	Weighted Avg
C3	0.90	0.82	0.86	0.87	0.86
C6	0.92	0.84	0.89	0.88	0.89
C4	0.90	0.88	0.89	0.89	0.89
C2	0.91	0.88	0.89	0.89	0.89

Tabella 22 - Risultati ottenuti nella classificazione dei picchi di consumo con dati aggregati a 15 minuti

Per orizzonte temporale pari a 4 si riportano i risultati per i diversi cluster:

1. Cluster 1:

	RMSE	MAE	SMAPE
C3	206.920459	123.922643	58.374881

Tabella 23- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 1 con orizzonte 4

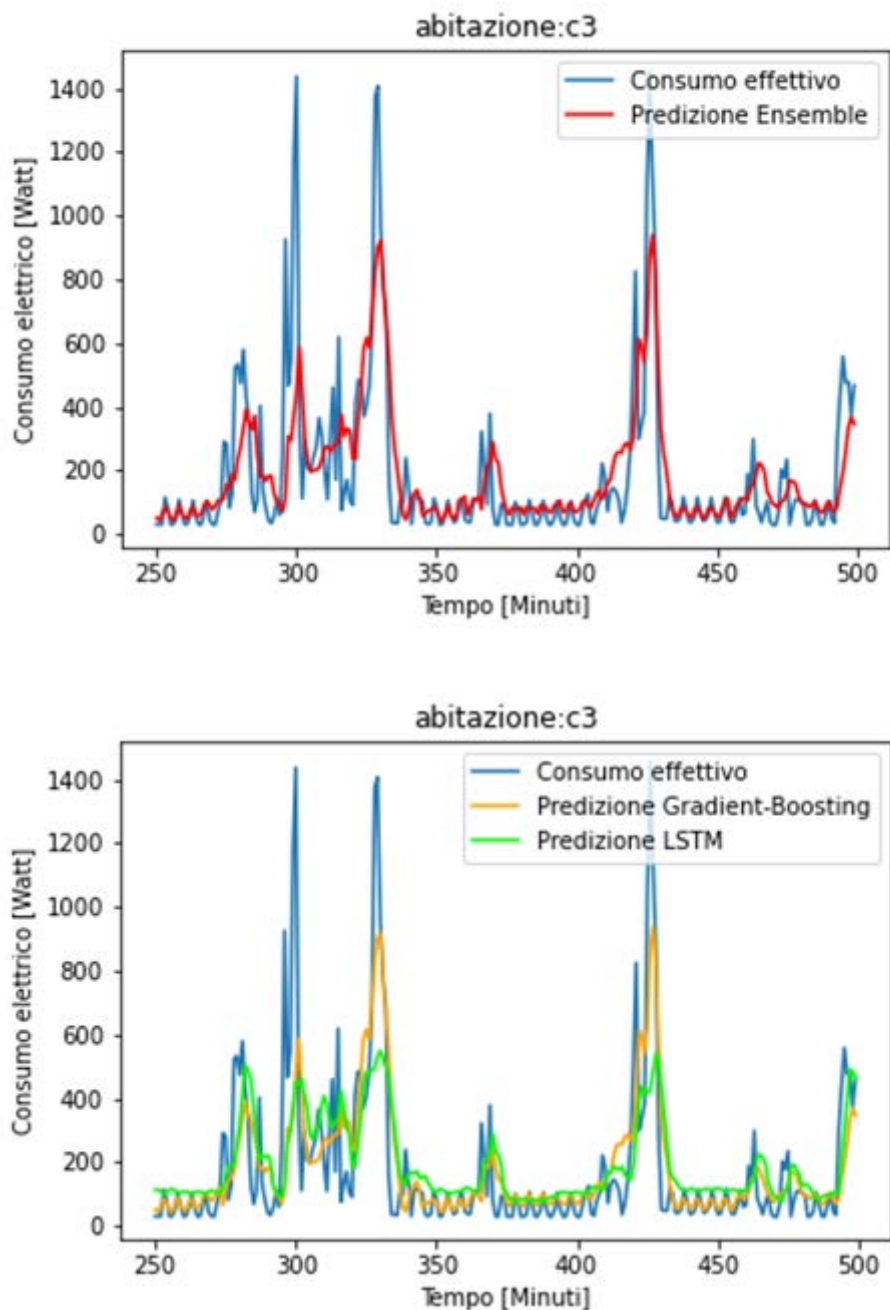


Figura 36: plot della previsione sul test set per abitazione C3

2. Per il cluster 2

	RMSE	MAE	SMAPE
C6	248.322161	111.235503	46.0924605
C1	209.465493	120.23336	34.706093

C5	276.599334	149.329858	44.4116094
C7	316.399566	133.32096	47.786926

Tabella 24- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 2 con orizzonte 4

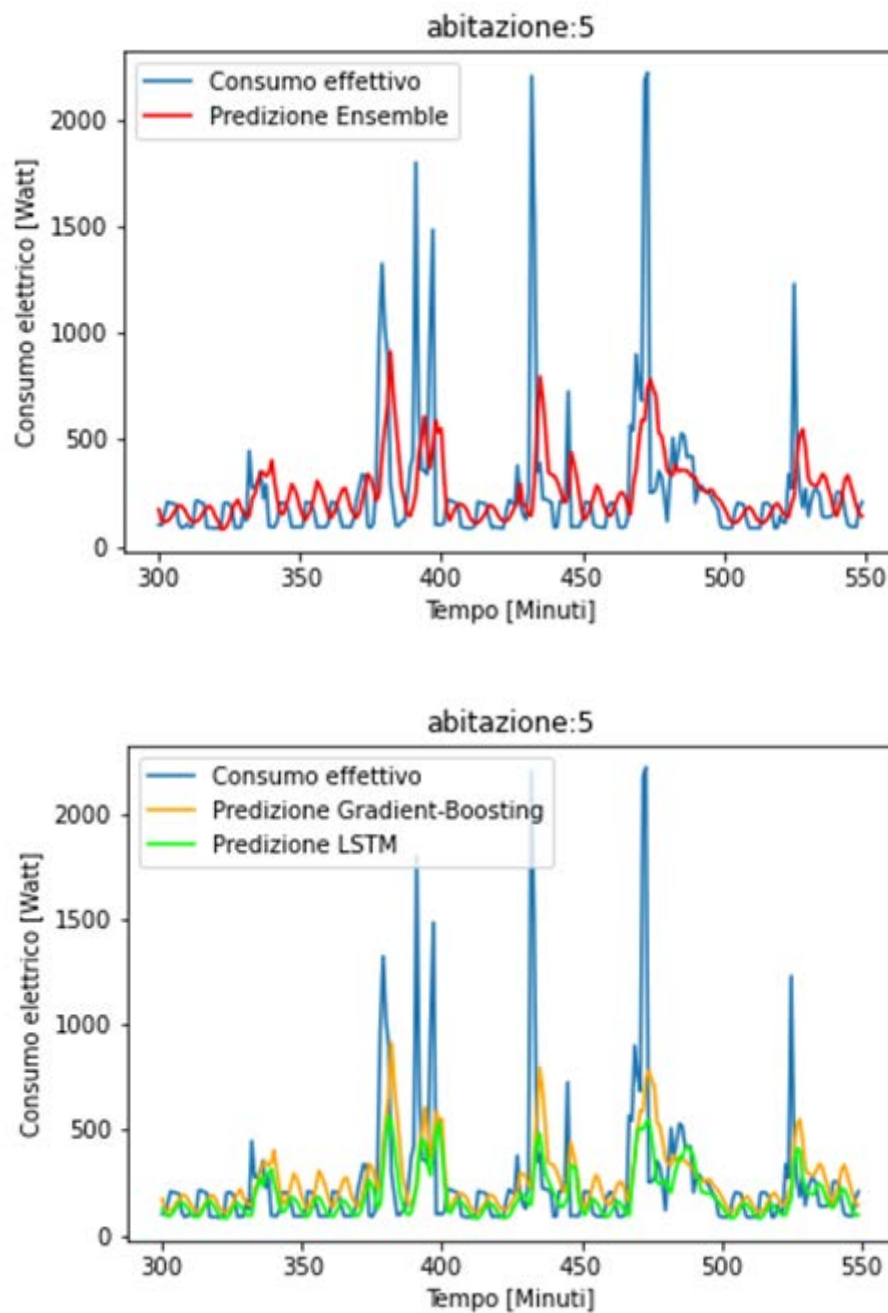


Figura 37: plot della previsione sul test set per abitazione C5

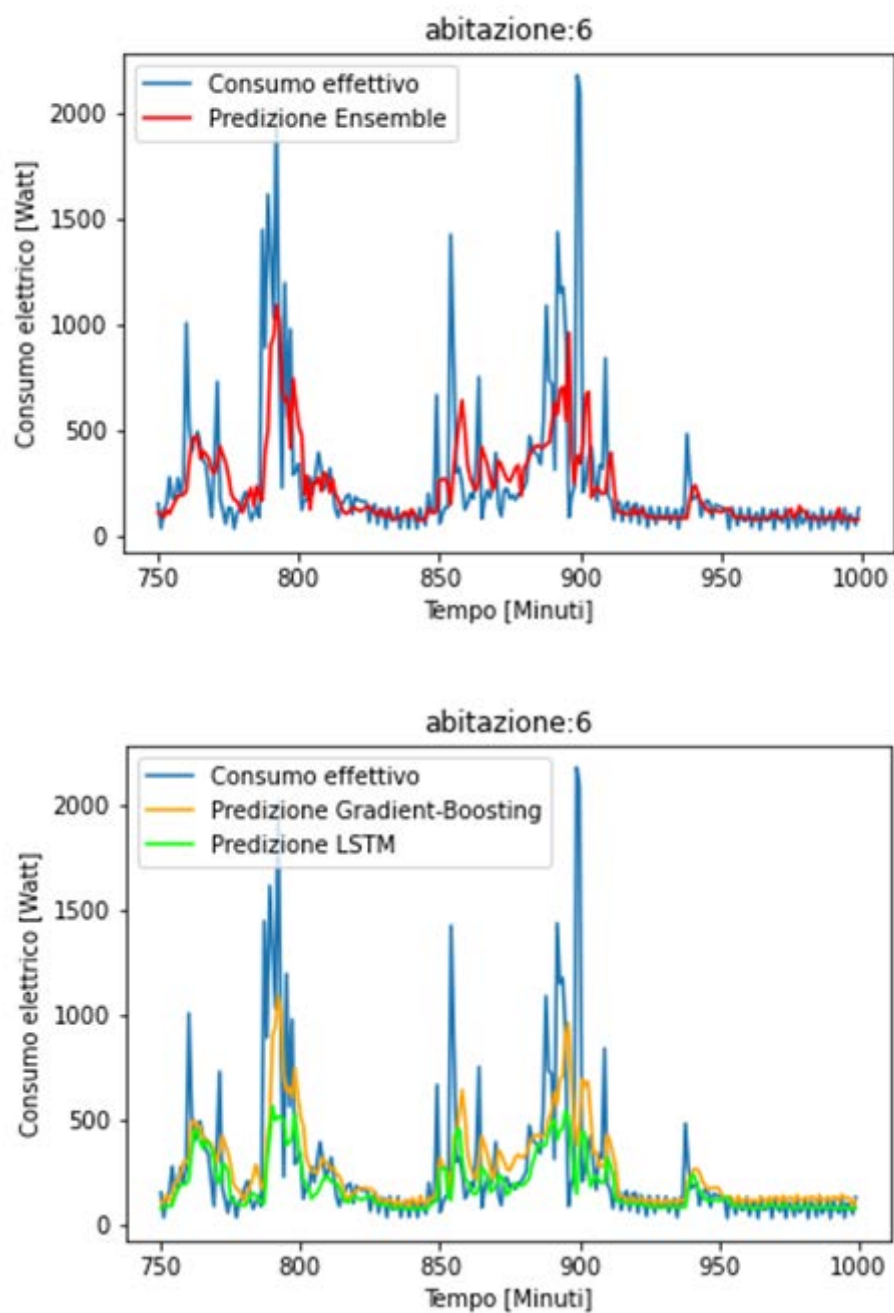


Figura 38: plot della previsione sul test set per abitazione C6

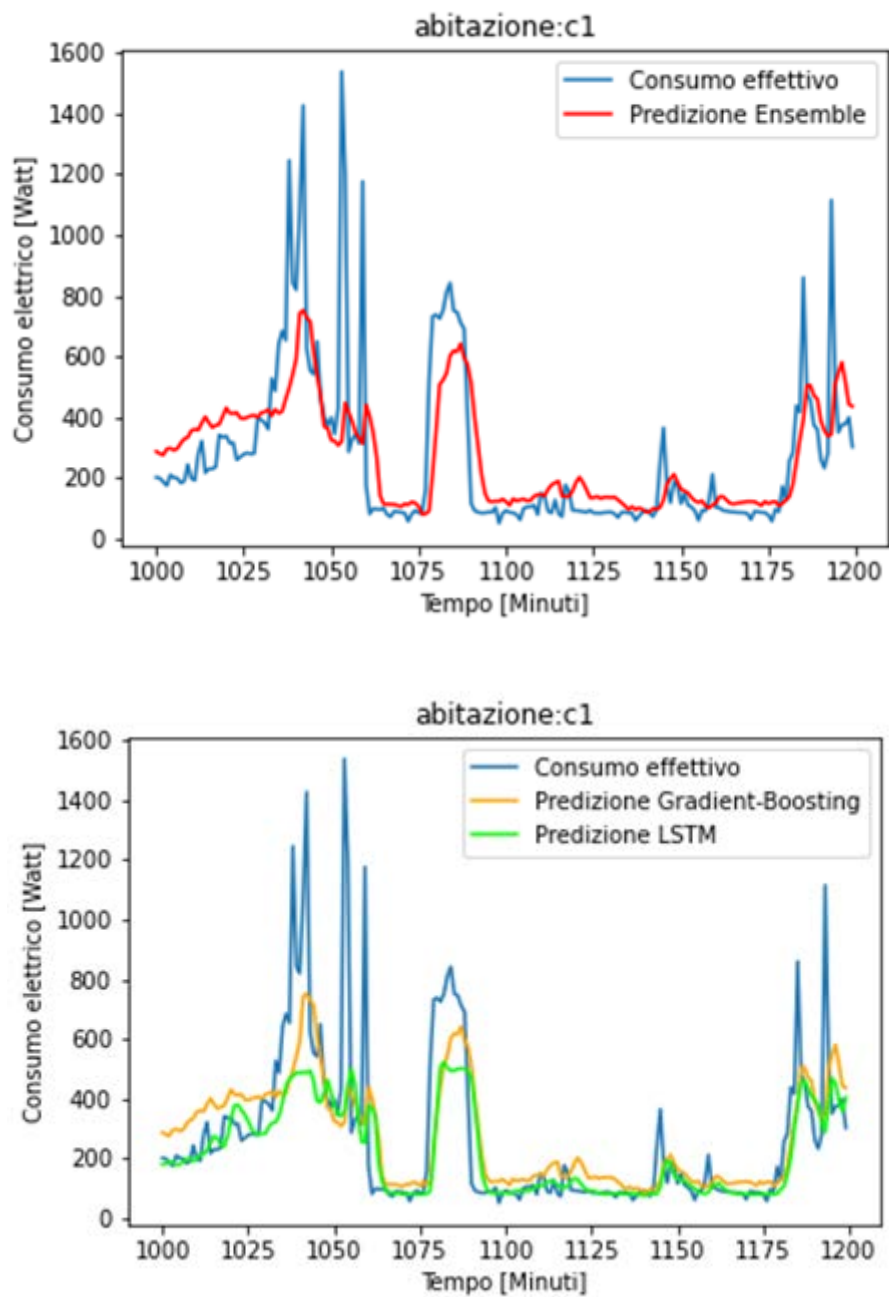


Figura 39: plot della previsione sul test set per abitazione C1

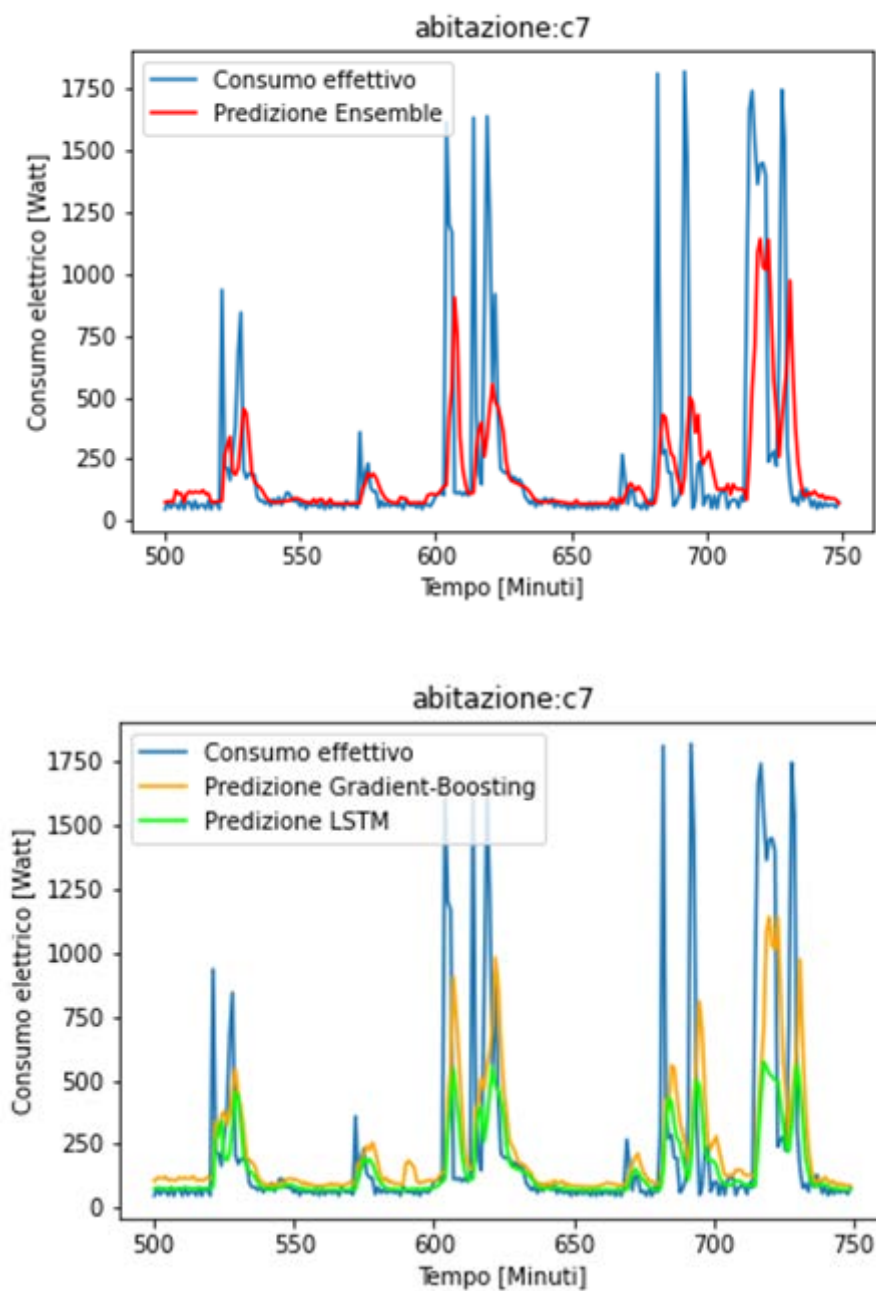


Figura 40: plot della previsione sul test set per abitazione C7

3. Per il cluster 3:

	RMSE	MAE	SMAPE
C4	195.647488	84.857168	53.448978

Tabella 25- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 3 con orizzonte 4

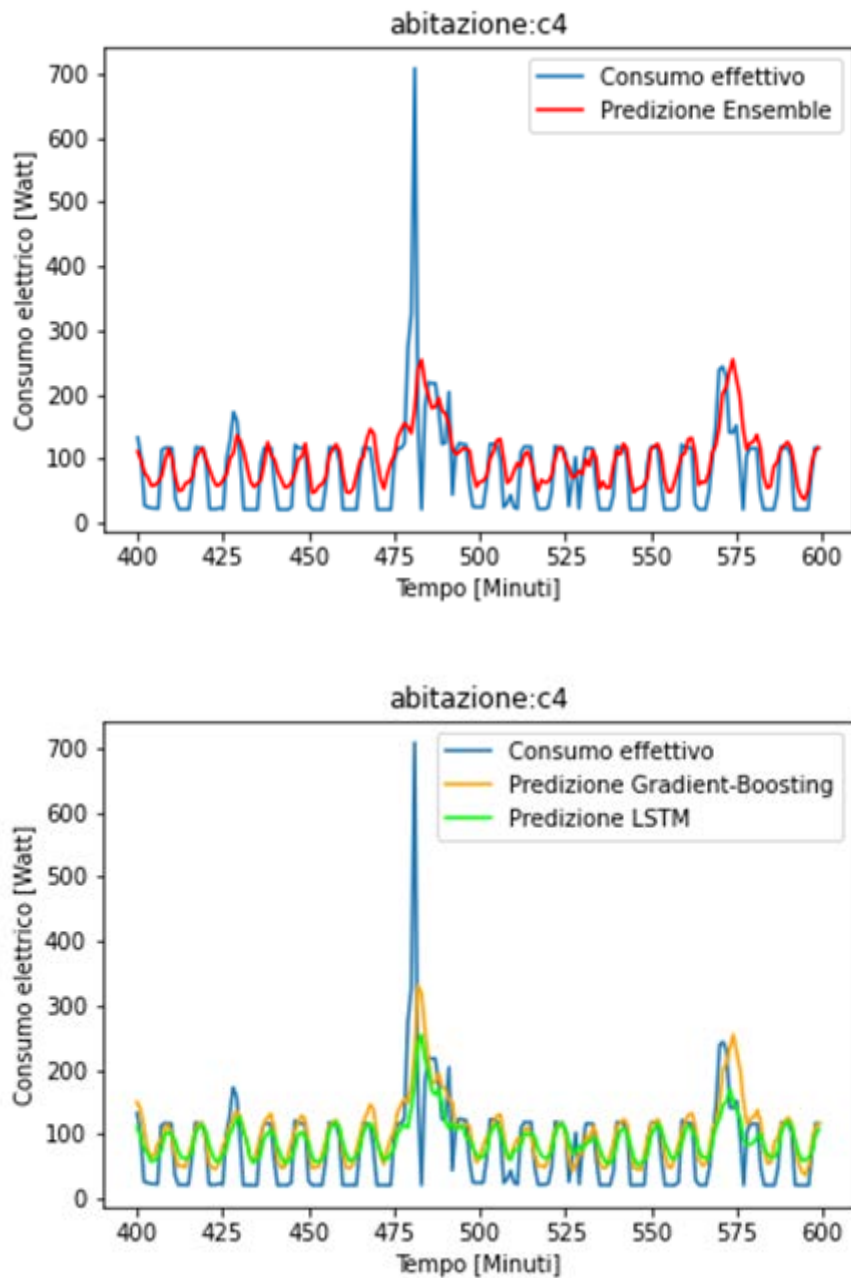


Figura 41: plot della previsione sul test set per abitazione C4

4. Per il cluster 4:

	RMSE	MAE	SMAPE
C2	232.371721	93.783312	56.784676

C8	159.575063	72.458010	56.649145
----	------------	-----------	-----------

Tabella 26- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 4 con orizzonte 4

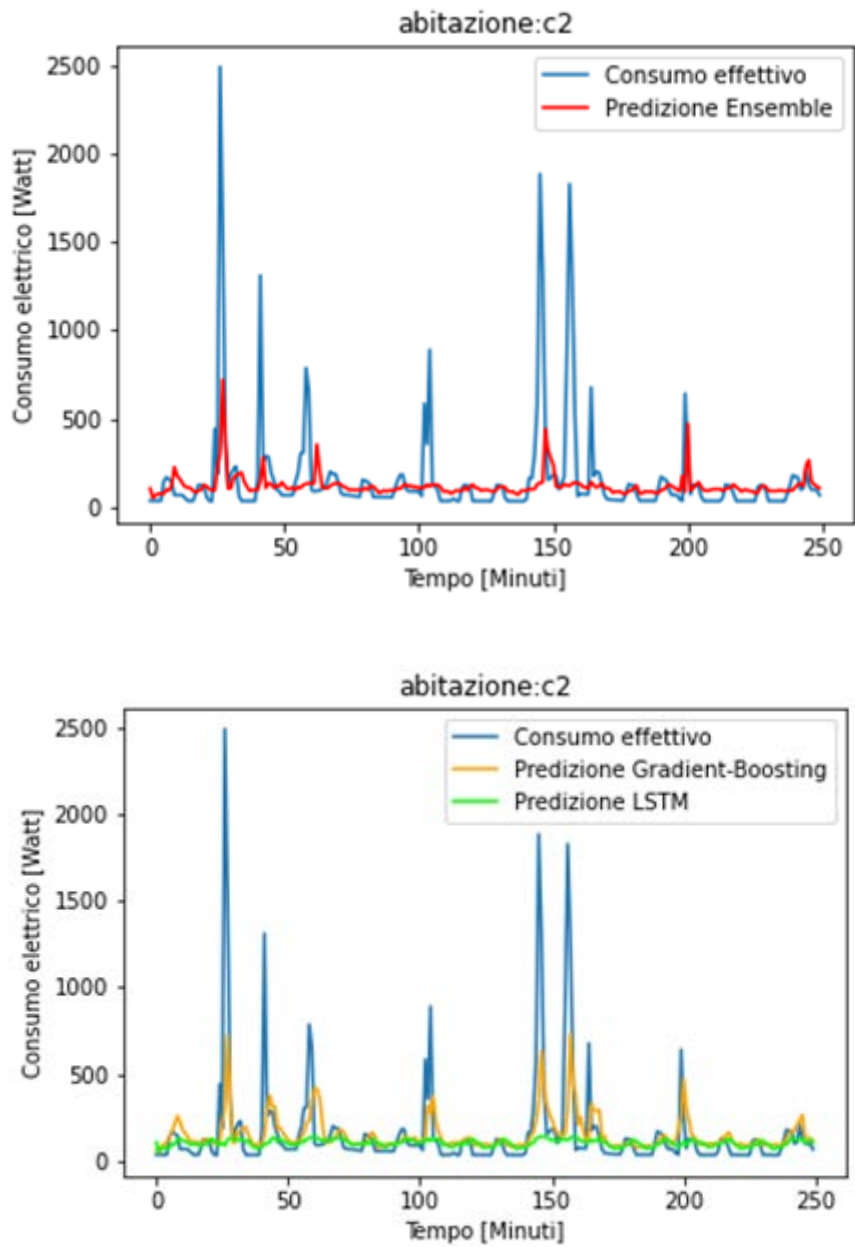


Figura 42: plot della previsione sul test set per abitazione C2

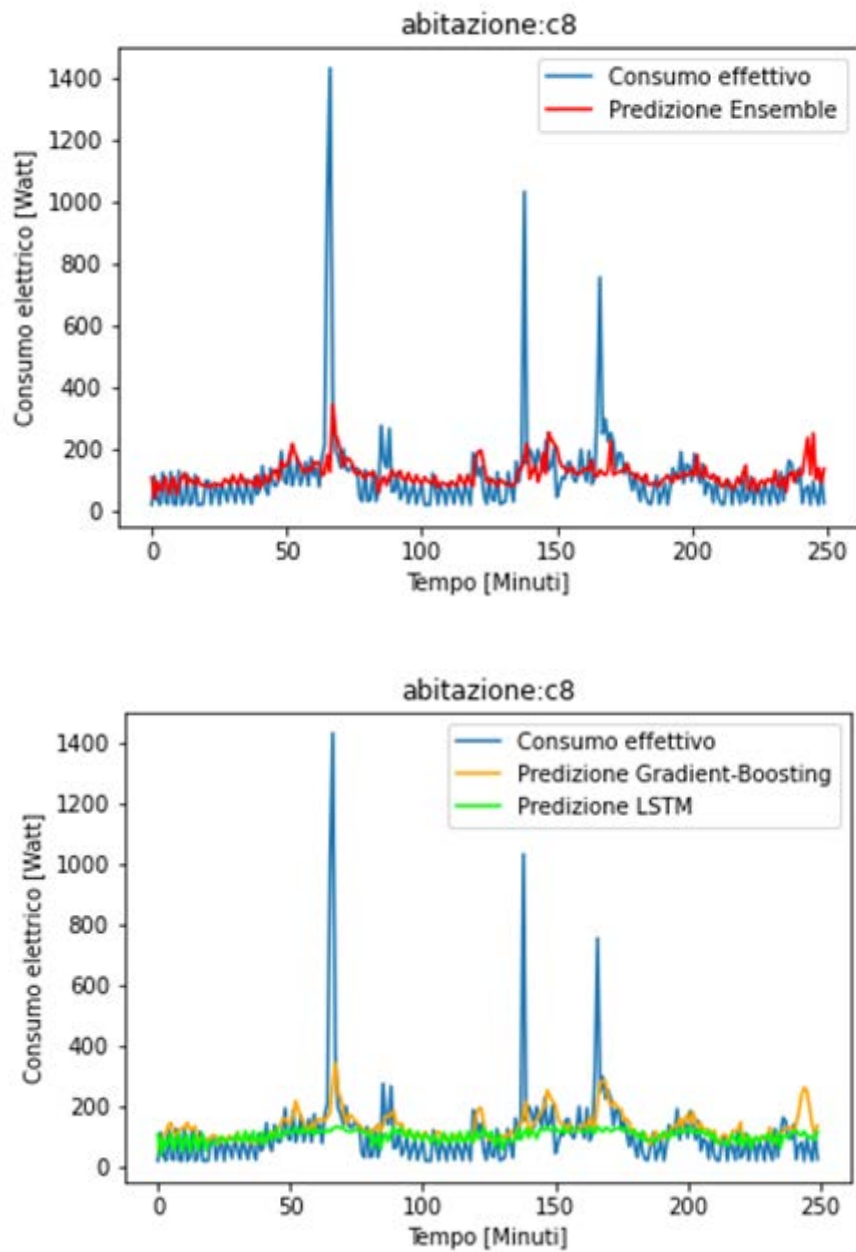


Figura 43: plot della previsione sul test set per abitazione C8

Con orizzonte temporale 6 si ottengono i seguenti risultati:

1. Cluster 1

	RMSE	MAE	SMAPE
C3	222.332627	135.511009	62.504583

Tabella 27- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 1 con orizzonte 6

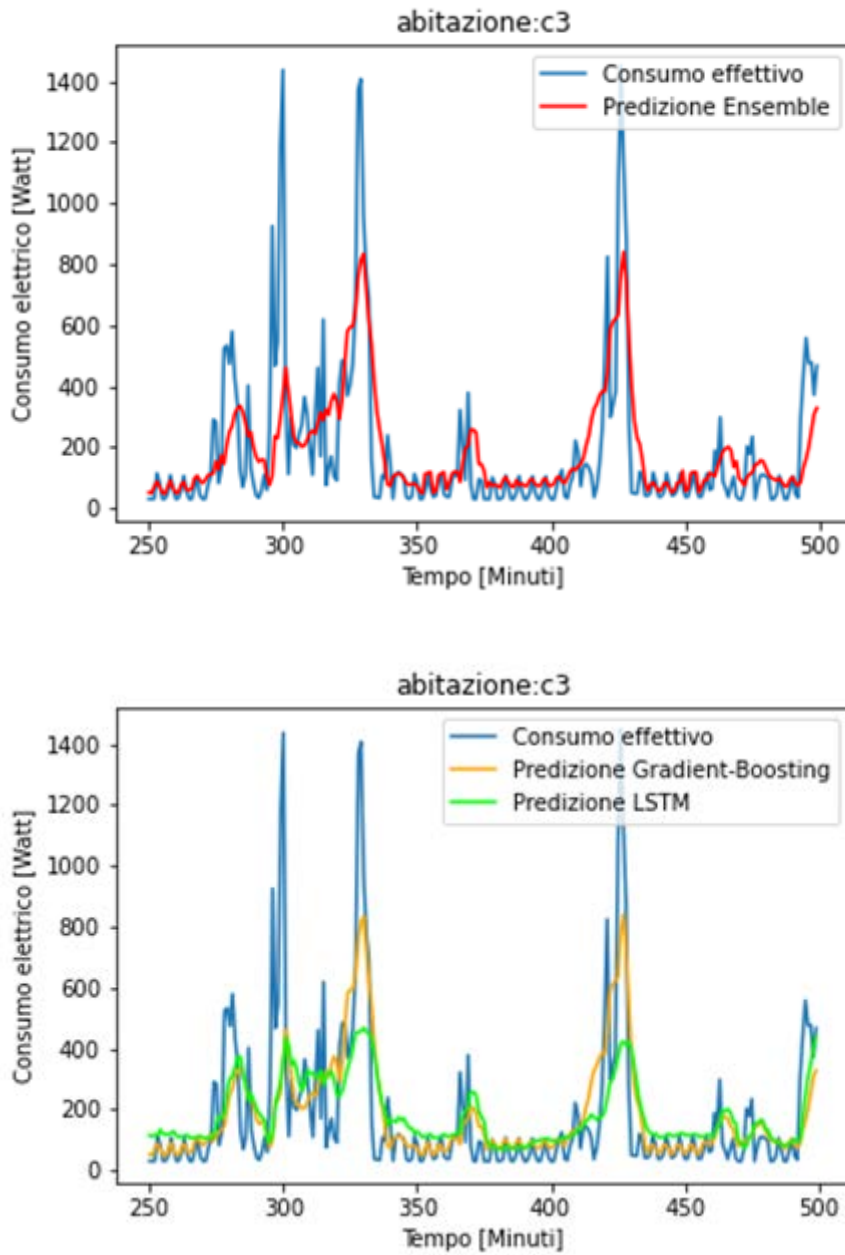


Figura 44: plot della previsione sul test set per abitazione C3

2. Per il cluster 2

	RMSE	MAE	SMAPE
C6	265.976762	122.941056	51.258226
C1	224.852023	133.699535	38.811911
C5	282.637876	154.491506	46.023432

C7	321.800509	149.821232	59.872307
----	------------	------------	-----------

Tabella 28- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 2 con orizzonte 6

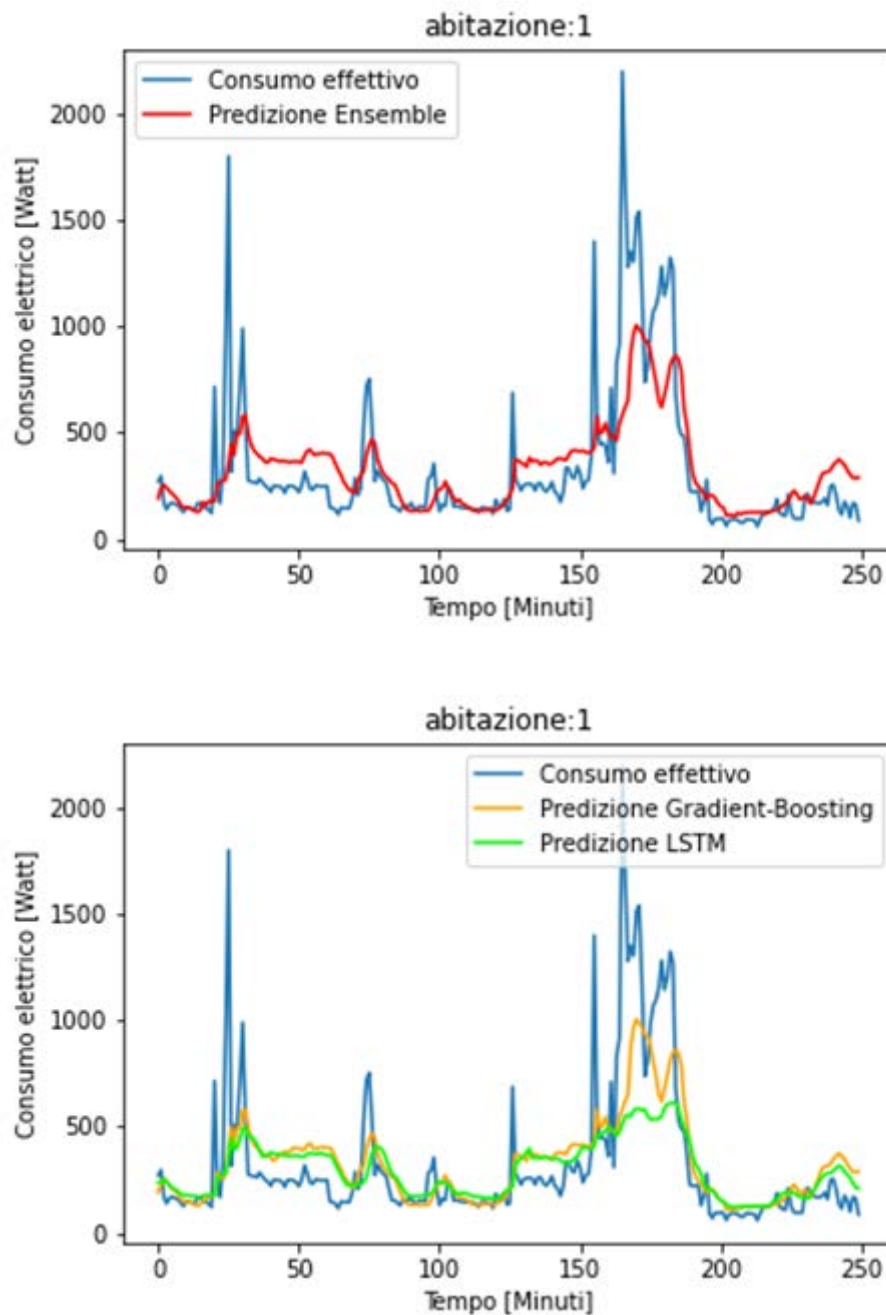


Figura 45: plot della previsione sul test set per abitazione C1

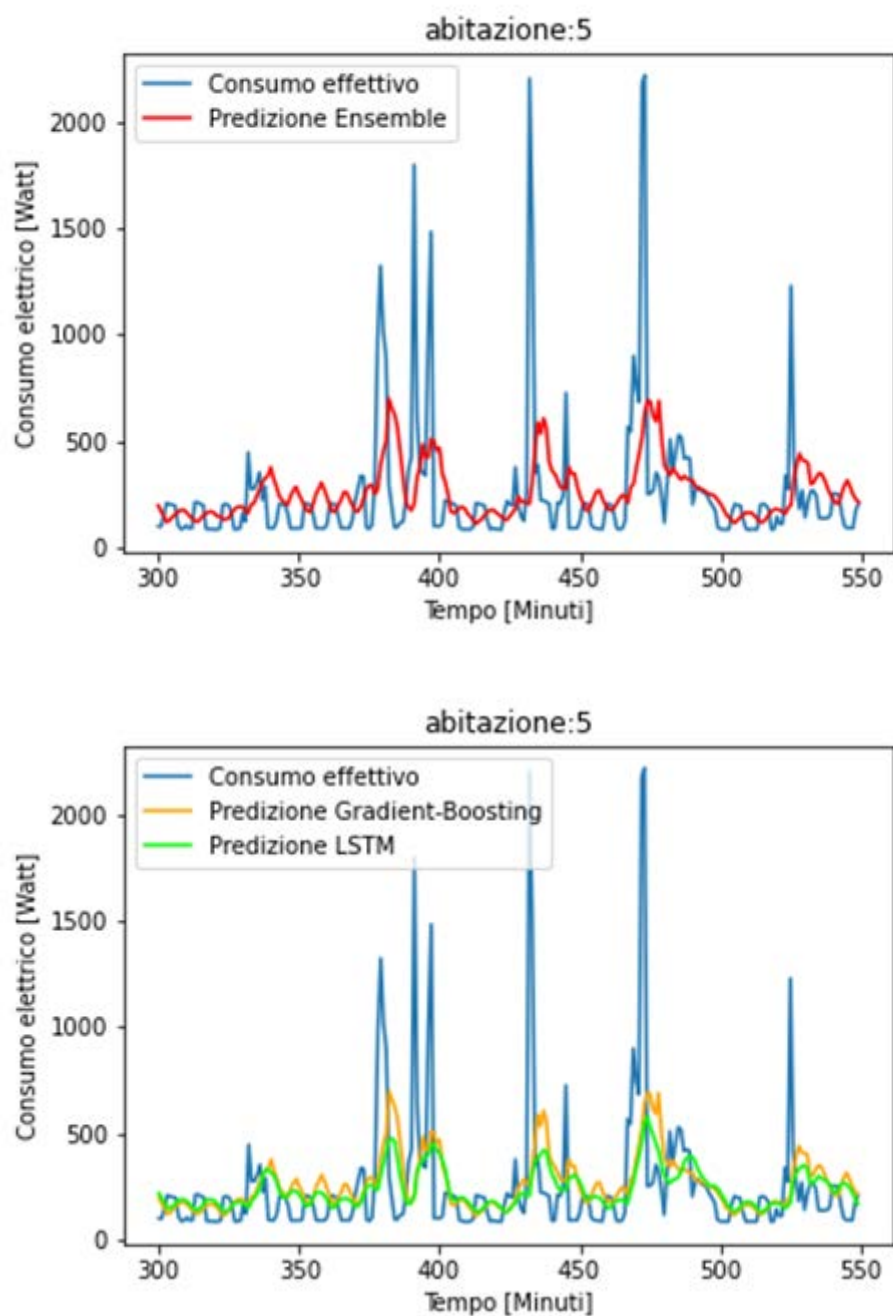


Figura 46: plot della previsione sul test set per abitazione C5

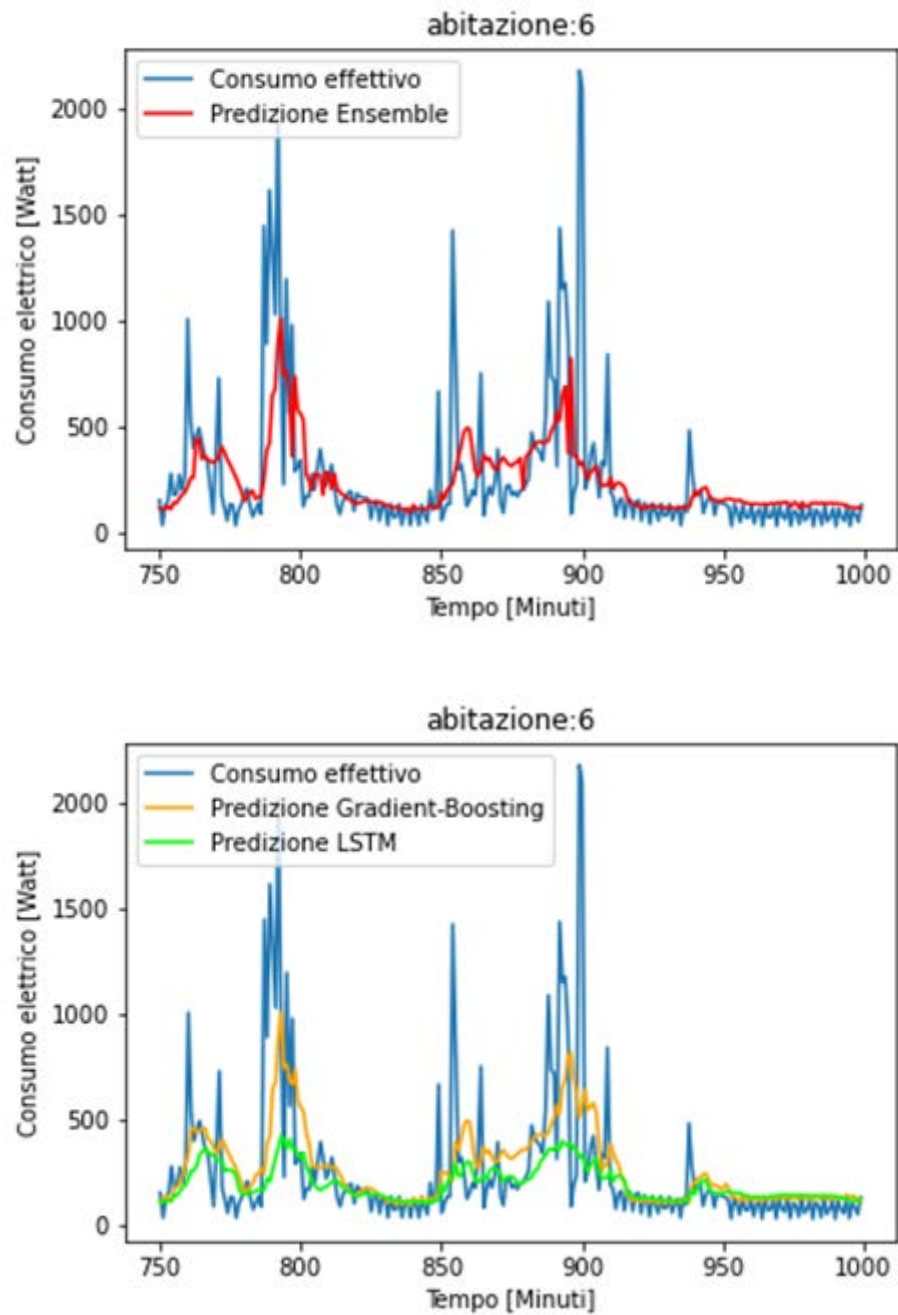


Figura 47: plot della previsione sul test set per abitazione C6

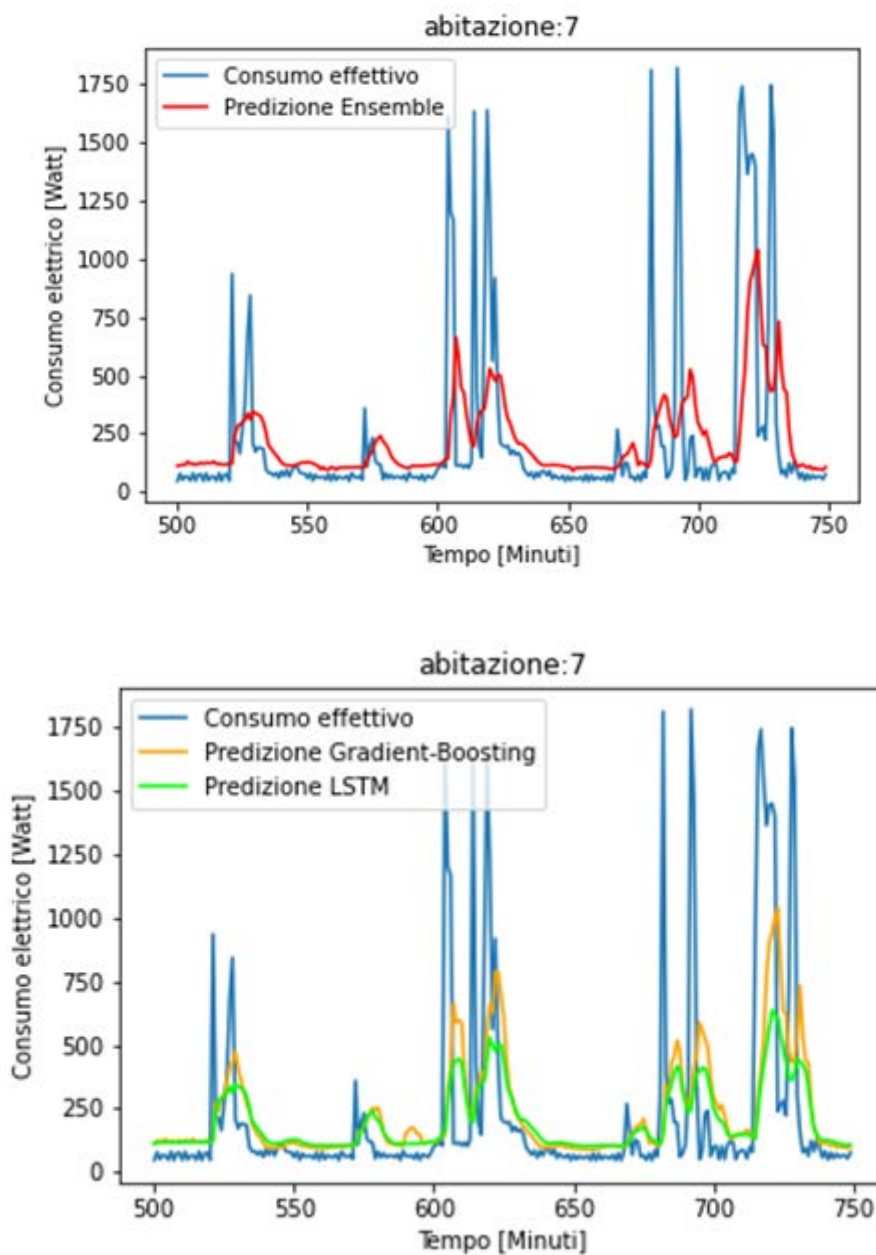


Figura 48: plot della previsione sul test set per abitazione C7

3. Cluster 3

	RMSE	MAE	SMAPE
C4	201.763803	85.772590	54.604813

Tabella 29- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 3 con orizzonte 6

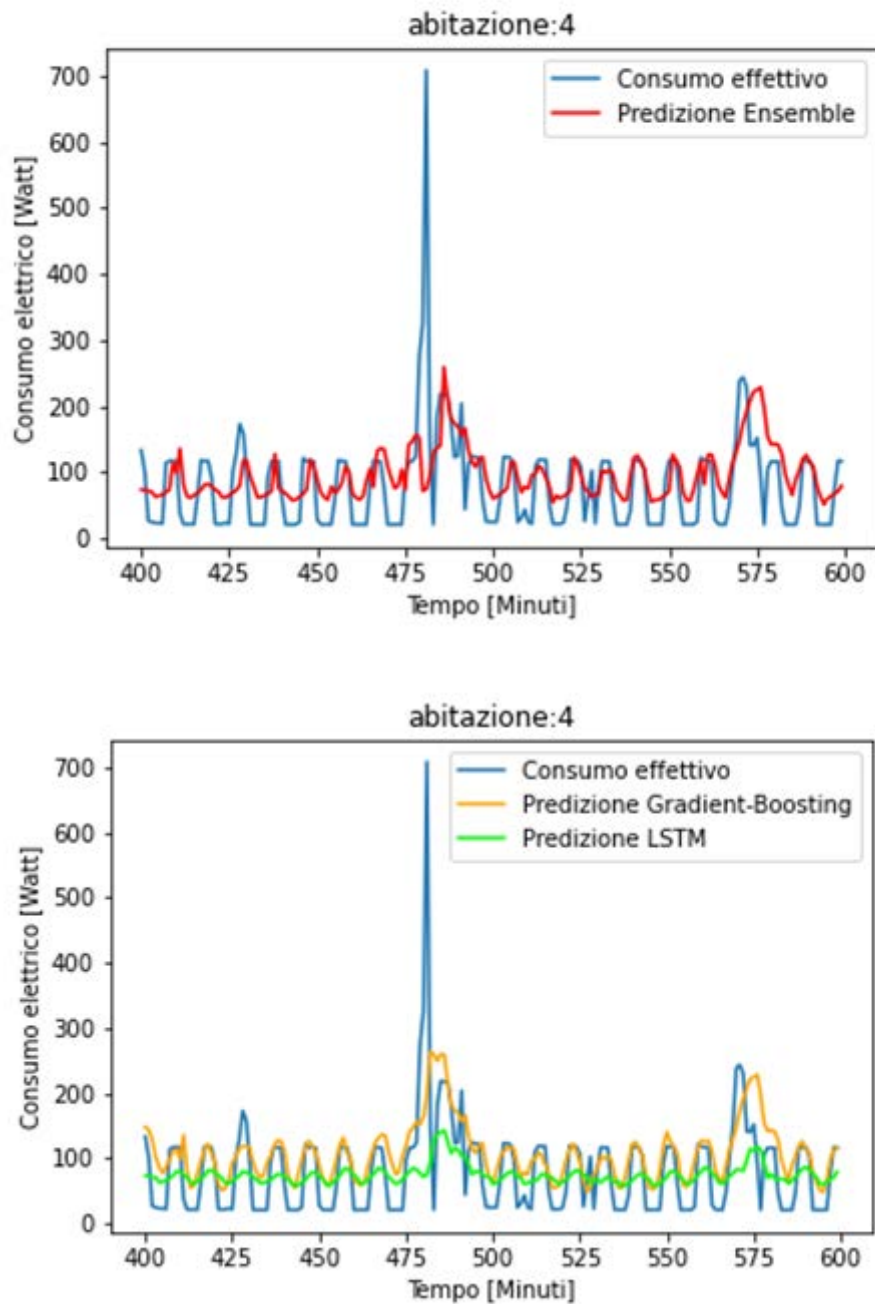


Figura 49: plot della previsione sul test set per abitazione C4

4. Per il Cluster 4

	RMSE	MAE	SMAPE
C2	237.837305	97.773238	59.433046
C8	161.827232	72.101547	56.221924

Tabella 30- Risultati ottenuti nella previsione con dati aggregati a 15 minuti per il cluster 4 con orizzonte 6

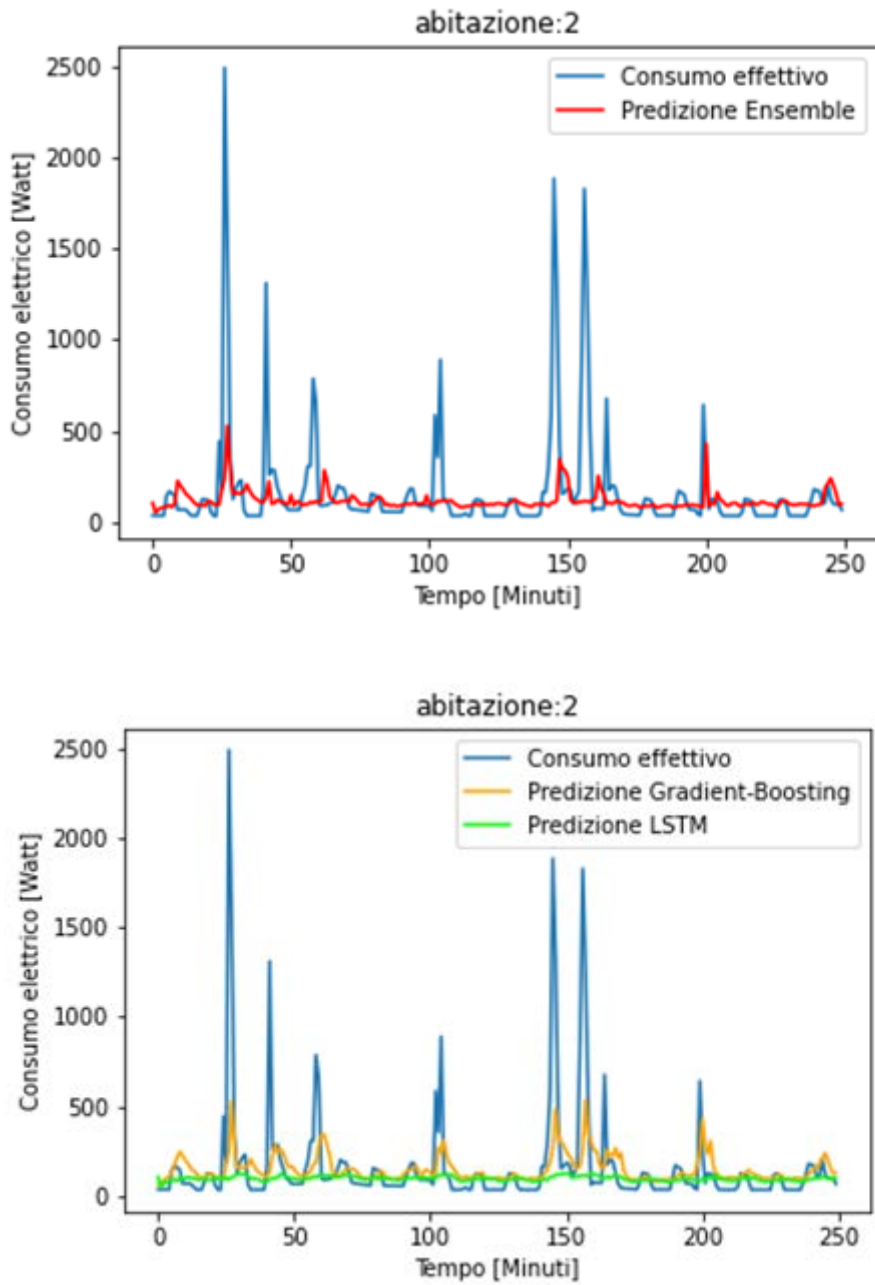


Figura 50: plot della previsione sul test set per abitazione C2

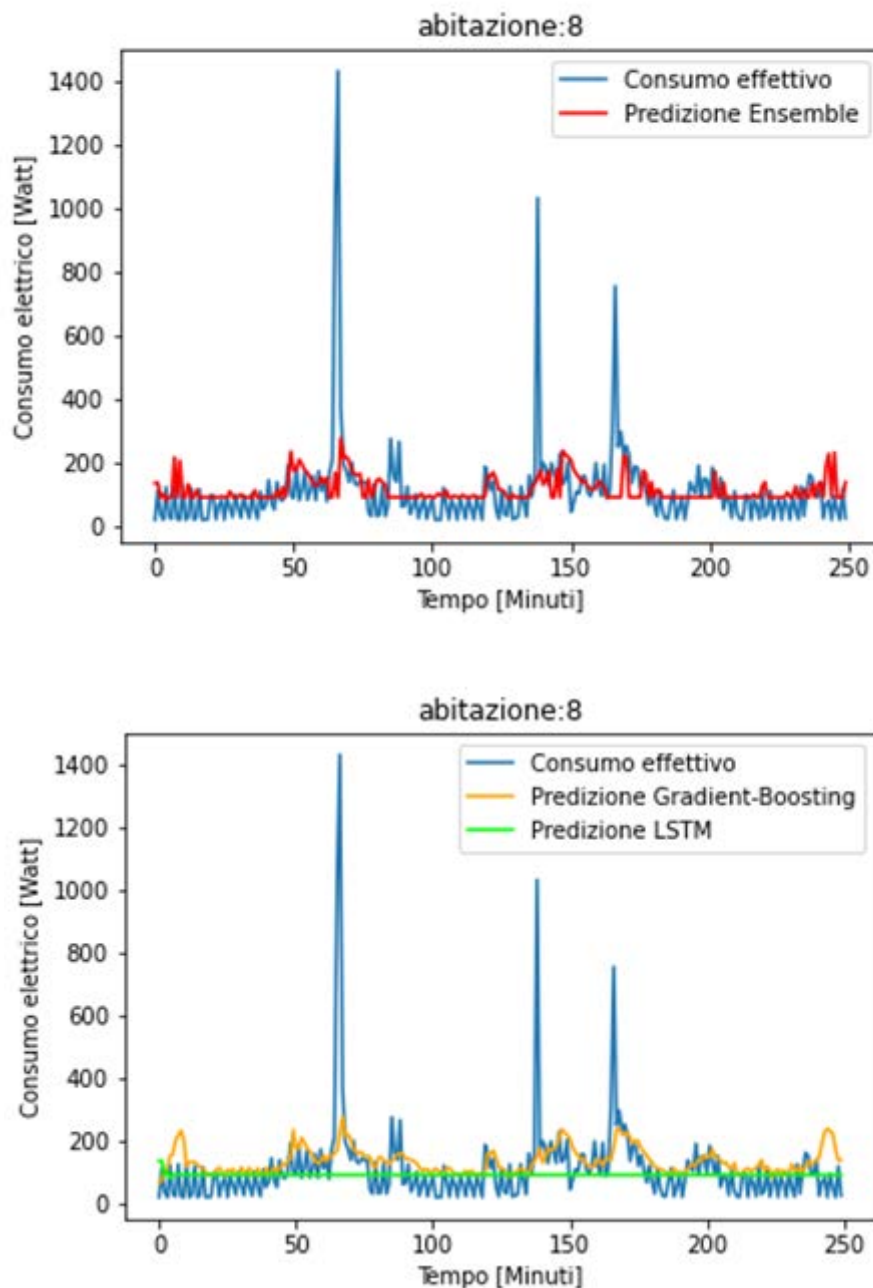


Figura 51: plot della previsione sul test set per abitazione C8

3.4.4 Confronto tra previsioni LSTM ed Ensemble Learning

In generale, per i consumi aggregati a 15 minuti, valgono le stesse conclusioni indicate per quelli aggregati a 5 minuti. L'introduzione dell'ensemble learning ha come conseguenza positiva il miglioramento generale dell'indice rmse. In alcuni casi si registra una diminuzione del mae e del mape ma, anche in questo caso, tale degradazione risulta essere poco significativa rispetto al buon impatto sull'rmse.

Abitazioni	Metrica	<u>Orizzonte 4</u>	<u>Orizzonte 4</u>	<u>Orizzonte 6</u>	<u>Orizzonte 6</u>
		<u>Ensemble</u>	LSTM	<u>Ensemble</u>	LSTM
C3	RMSE	206.9204591	218.6579367	222.3326273	231.5249609
	MAE	123.922643	131.844707	135.5110093	137.2918739
	SMAPE	58.37488142	65.81641137	62.50458329	67.26479176
C6	RMSE	248.3221608	255.8468535	265.9767624	260.6264531
	MAE	111.2355025	104.0688893	122.9410565	122.5901525
	SMAPE	46.09246054	44.46955372	51.2582262	52.87224268
C1	RMSE	209.4654935	232.9616048	224.8520228	232.6662039
	MAE	120.233358	114.0010532	133.6995355	135.0189573
	SMAPE	34.7060932	30.48277468	38.81191066	39.37112175
C5	RMSE	276.5993336	295.9632875	282.6378758	289.8450388
	MAE	149.3298577	135.5529464	154.4915063	149.0390756
	SMAPE	44.41160941	39.8223983	46.02343194	44.5190946
C9	RMSE	316.3995655	318.3586702	321.8005086	322.0905586

C7	MAE	133.320957	114.4557146	149.8212317	144.8808959
	SMAPE	47.78692595	37.32335292	59.87230683	60.67845982
C4	RMSE	195.6474881	200.2730618	201.7638029	221.4253797
	MAE	84.85716794	84.24166464	85.77259007	100.1333148
	SMAPE	53.44897839	54.59086062	54.60481322	63.42167274
C2	RMSE	232.3717205	238.3707974	237.837305	240.8232051
	MAE	93.78331179	87.87174655	97.7732384	96.12995867
	SMAPE	56.78467648	52.98137815	59.43304574	59.91859873
C8	RMSE	159.5750634	159.3801505	161.8272319	162.8239144
	MAE	72.45800984	63.94564904	72.1015472	62.78238694
	SMAPE	56.64914502	52.90463635	56.22192368	52.13347739

Tabella 31- Confronto tra ensemble learning e LSTM per dati aggregati a 15 minuti

4 Analisi dei dati delle nuove case forniti a novembre 2021

A novembre 2021 ci sono stati forniti nell'ambito del progetto dei nuovi dati, relativi a 10 abitazioni diverse per cui si hanno a disposizione 5 mesi di dati scarsi da giugno 2021 a fine ottobre 2021, con un numero medio elevato di plug per ogni abitazione, che quindi sono molto interessanti dal punto di vista della disaggregazione. Di seguito sono riportate le plug per ogni casa

- Nella casa H1 sono presenti i plug su Asciugatrice, Computer dei ragazzi, Computer dello studio, Condizionatore della camera matrimoniale, Condizionatore della camera dei ragazzi,

Condizionatore del soggiorno, Illuminazione, Lavastoviglie, Lavatrice, Scaldabagno e pompa di calore, Tv camera matrimoniale, Tv soggiorno

- nella casa H2 sono presenti i plug su Asciugatrice, Condizionatore del soggiorno, plug Condizionatore della zona notte, Lavastoviglie, Lavatrice, Scaldabagno e pompa di calore. Purtroppo però i consumi complessivi per questa casa sono disponibili per un periodo temporale inferiore a 3 settimane, che rende la casa inutilizzabile.
- nella casa H3 si hanno i plug su Lavastoviglie, Lavatrice e Tv
- nella casa H4 si hanno i plug su condizionatore della camera da letto, Frigorifero, Lavastoviglie e Lavatrice.
- nella casa H5 si hanno le plug su Frigorifero, Lavatrice, Macchina del caffè, Tv. La casa non è però utilizzabile perchè nel main è presente circa l'87% di sommerso, ovvero consumi non monitorati
- nella casa H6 sono presenti i plug su Asciugatrice, Condizionatore della camera da letto, Frigorifero, Lavastoviglie, Lavatrice
- nella casa H9 sono presenti i plug su Asciugatrice, Condizionatore della camera letto, Condizionatore della seconda camera da letto, Lavastoviglie, Lavatrice.
- nella casa H10 sono presenti i plug su Frigorifero, Lavastoviglie, Lavatrice; purtroppo questa casa è inutilizzabile in quanto è presente meno di un mese di dati
- nella casa H11 sono presenti i plug su Condizionatore in camera da letto, Condizionatore in sala hobby, Frigo, Lavastoviglie, Lavatrice
- nella casa H12 sono presenti i plug Frigo, Lavastoviglie, Lavatrice: anche questa casa non è utilizzabile perchè nel main è presente circa il 73% di sommerso.

Questo nuovo dataset presenta diverse criticità: dal punto di vista della disaggregazione dei consumi, per applicare l'approccio tramite rete deep serve una quantità di dati maggiori, tenuto anche conto che uno dei mesi disponibili è quello di agosto, in cui i consumi sono quasi nulli e ci sono pochissime attivazioni di elettrodomestici. Dal punto di vista invece della previsione del main, non abbiamo a disposizione nessun tipo di informazione simile a quella fornita nei questionari, e quindi non possiamo replicare l'approccio descritto nella sezione precedente.

Tuttavia, recentemente in [21] è stata proposta una nuova tecnica di disaggregazione di consumi basata sull'ottimizzazione che può essere applicata anche in presenza di quantità di dati molto ridotte, e permette di disaggregare contemporaneamente il consumo di un numero elevato di elettrodomestici. In particolare, il problema della disaggregazione dei consumi energetici viene formulato come un problema di ottimizzazione quadratico con vincoli lineari e variabili binarie (BQP). In questa formulazione si assume che la potenza elettrica complessiva richiesta dall'abitazione $y(t)$, in ogni istante di tempo t , sia disponibile e possa essere vista come

$$y(t) = \sum_{i=1}^N y_i(t) + \epsilon(t)$$

dove N è il numero di dispositivi che si vuole monitorare, $y_i(t)$ è la potenza elettrica che l'apparecchio i consuma all'istante temporale t e $\epsilon(t)$ rappresenta un termine di rumore che comprende, oltre ad eventuali apparecchi non monitorati, che ne sono la componente predominante, anche l'errore di misurazione. Di conseguenza per $t \in T$ il problema della disaggregazione del carico può essere banalmente formulato come un problema di minimizzazione dell'errore

$$\begin{aligned} \min \quad & \left(y(t) - \sum_{i=1}^N p_i^T x_i(t) \right)^2 \\ \text{s.t.} \quad & x_i(t) \in \{0, 1\}^{S_i} \quad \forall i \in \mathcal{N}, \end{aligned}$$

dove p_i è il vettore che contiene i livelli di potenza a cui opera il dispositivo i e $x_i(t) \in \{0, 1\}^{S_i}$ è un vettore di variabili di stato dell'elettrodomestico i . La j -esima componente di $x_i(t)$ viene impostata ad 1 se lo stato j del dispositivo i è considerato attivo al tempo t , 0 altrimenti.

Prendiamo quindi come punto di partenza la formulazione non vincolata appena descritta e aggiungiamo sia termini di regolarizzazione che vincoli. I termini di regolarizzazione mirano a gestire quanto più possibile la presenza di dispositivi non monitorati nel carico totale, mentre i vincoli cercano di modellare al meglio le curve di consumo dei dispositivi monitorati.

La funzione obiettivo da minimizzare si compone di tre termini sommati tra loro: uno principale e due di penalità; questi ultimi due termini sono stati aggiunti sulla base della conoscenza pregressa del problema con il fine di migliorare le prestazioni del modello. Il termine principale della funzione obiettivo consiste nella somma su tutto l'orizzonte temporale T preso in considerazione dei quadrati delle differenze tra potenza elettrica complessiva e la somma delle componenti disaggregate:

$$f(x(t)) = \sum_{t=1}^T \left(y(t) - \sum_{i=1}^N \hat{y}_i(t) \right)^2$$

Sotto l'assunzione che ogni dispositivo lavora seguendo uno o più livelli di potenza, $f(x(t))$ può essere riscritta come segue:

$$f(x(t)) = \sum_{t=1}^T \left(y(t) - \sum_{i=1}^N (p_i^T x_i(t)) \right)^2$$

Al termine principale ne viene aggiunto un secondo che punta a rendere la richiesta di potenza di ogni apparecchio a tratti costante nel tempo; questo permette di evitare frequenti cambi di stato che risultano essere poco realistici per la maggior parte dei dispositivi. Per forzare questo comportamento si penalizzano le transizioni di stato degli elettrodomestici durante l'orizzonte temporale; ciò può essere eseguito applicando una tecnica nota in letteratura come group lasso secondo la quale si penalizza la norma della differenza tra $x_i(t)$ e $x_i(t-1)$, in questo ambito viene utilizzata la norma l_2 [23]. Possiamo quindi scrivere questo termine di penalità come segue

$$g_1(x(t)) = \lambda_1 \sum_{t=2}^T \sum_{i=1}^N w_i |x_i(t) - x_i(t-1)|_2^2$$

nella quale λ_1 è il parametro di penalità che rappresenta quanto questo termine deve influire sulla funzione obiettivo; invece il vettore $w \in \mathbb{R}_+^N$ contiene dei pesi che determinano quanto contribuiscono i singoli dispositivi al termine di penalità. Questi ultimi dovrebbero essere scelti in base alla predisposizione dei dispositivi a cambiamenti frequenti di stato: nello specifico il peso w_i dovrebbe essere impostato a un valore inversamente proporzionale alla frequenza di questi

cambiamenti in modo da penalizzare maggiormente una transizione di stato in quei dispositivi che tendono ad avere un comportamento più costante nel tempo.

L'ultimo termine che compone la funzione obiettivo mira a rendere il modello robusto al rumore e con un maggior grado di sparsità preferendo tra le varie combinazioni quella con il minor numero di dispositivi attivi. Così facendo si evita, qualora il rumore sia significativo, che tutti i dispositivi siano messi in uno stato attivo per la maggior parte del tempo. Per raggiungere questo scopo si utilizza la norma l_2 del vettore binario $x_i(t)$ come segue

$$g_2(x(t)) = \lambda_2 \sum_{t=2}^T \sum_{i=1}^N l_i (1 - s_i(t)) \|x_i(t)\|_2,$$

dove λ_2 è il parametro di penalità che va a misurare l'impatto che il termine di penalità deve avere. Infine il termine $(1 - s_i(t))$ permette di penalizzare maggiormente l'attivazione dei dispositivi in momenti dell'orizzonte temporale in cui tipicamente non vengono utilizzati (ad esempio l'utilizzo del forno elettrico in piena notte).

La prima classe di vincoli che è stata inserita nella formulazione riguarda il numero di stati contemporaneamente attivi per lo stesso dispositivo; infatti è ragionevole assumere che ogni elettrodomestico possa trovarsi al più in una modalità operativa per volta

$$1_{S_i}^T x_i(t) \leq 1 \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N}$$

Ci sono, inoltre, alcuni dispositivi detti always on che si trovano in uno stato attivo tutto il tempo, per questi il vincolo dovrebbe essere rispettato all'uguaglianza:

$$1_{S_i}^T x_i(t) = 1 \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{O}$$

dove l'insieme \mathcal{O} contiene gli indici i di tutti i suddetti elettrodomestici. Andando ad analizzare i profili energetici di alcuni dispositivi si può notare come questi tendano ad avere sempre lo stesso tipo di variazioni di stati, questa informazione può quindi essere sfruttata per disambiguare maggiormente dispositivi che potrebbero avere livelli di potenza simili. Per migliorare l'accuratezza del modello e permettergli di distinguere più facilmente i dispositivi caratterizzati da molteplici stati sono state introdotte due classi di variabili binarie ausiliare $u_i(t), d_i(t) \in \{0,1\}^{S_i}$ che vanno a rappresentare le transizioni verso l'alto (ossia quando un dispositivo passa da essere spento ad avere uno degli stati operativi attivo) e quelle verso il basso (quando un dispositivo, contrariamente a quanto avveniva nelle transizioni verso l'alto, passa da uno stato attivo a essere spento).

$$\begin{aligned} x_i(t) - x_i(t-1) &= u_i(t) - d_i(t) \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \\ u_i(t) + d_i(t) &\leq 1_{S_i}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \end{aligned}$$

Grazie a questi vincoli possiamo vedere che, se il dispositivo i cambia modalità operativa al tempo t , allora le componenti di $x_i(t) - x_i(t-1)$ potranno assumere i valori: $-1, 0$ e 1 . In particolare se la componente j -esima è settata a 1 il vincolo forza la j -esima componente del vettore $u_i(t)$ a 1 rappresentando così una transizione verso l'alto; se invece è settata a -1 viene forzata a 1 la j -esima componente del vettore $d_i(t)$ andando a modellare una transizione verso il basso. Invece il secondo vincolo permette di evitare che le componenti di $u_i(t)$ e $d_i(t)$ siano contemporaneamente impostate ad 1 . Le variabili ausiliarie appena introdotte sono necessarie sia per modellare i vincoli che andranno a far rispettare il comportamento basato su macchine a stati sia per modellare altri vincoli che puntano a rendere più accurato il modello finale andando a sfruttare la conoscenza pregressa riguardo ai dispositivi da monitorare. Al fine di permettere al modello di identificare

quanto più accuratamente possibile i dispositivi che lavorano come macchine a stati sono stati aggiunti, come accennato in precedenza, i vincoli che regolano proprio le transizioni di stato:

$$U_i u_i(t) = D_i d_i(t), \forall t \in \mathcal{T}, i \in \mathfrak{S}$$

dove $\mathfrak{S} \subseteq \mathcal{N}$ è l'insieme che contiene gli indici degli apparecchi caratterizzati da un comportamento basato su macchine a stati, $U_i, D_i \in \{0,1\}^{S_i \times S_i}$ sono matrici binarie e sparse che vanno a definire l'ordine in cui gli stati devono essere attraversati in un ciclo di attivazione. modo che il modello rispetti il corretto funzionamento della macchina a stati.

Se si analizzano i profili energetici dei vari dispositivi ci si rende inoltre conto che sarebbe irragionevole permettere che un apparecchio cambi stato in ogni istante di tempo, anzi tipicamente ogni stato è caratterizzato da un tempo minimo di permanenza che a seconda del dispositivo e della modalità operativa può essere più o meno lungo. Per permettere al modello di far rispettare questo comportamento è stata aggiunta la classe di vincoli:

$$\sum_{\tau=t}^{t+a_i^j-1} x_i^j(\tau) \geq a_i^j u_i^j(t) \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \forall j \in S_i$$

dove il vettore di parametri a_i^j contiene i minimi tempi di permanenza negli stati dell'elettrodomestico i . Analogamente, è anche logico vincolare il massimo tempo di permanenza in uno stato; questo permette di rendere non ammissibili tutte quelle soluzioni in cui uno stato viene considerato attivo continuamente, situazione che, in presenza di un quantitativo di rumore elevato, si potrebbe presentare frequentemente. Proprio per evitare che il modello scelga un comportamento erraneo come quello appena descritto è stata introdotta un'ulteriore classe di vincoli

$$\sum_{\tau=t}^{t+b_i^j} x_i^j(\tau) \leq b_i^j \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \forall j \in S_i$$

dove il vettore di parametri b_i^j è il corrispettivo di a_i^j per i massimi tempi di permanenza negli stati.

Per limitare la presenza di false attivazioni, molto comuni quando si ha a che fare con dati in cui il rumore è una componente significativa, è stato aggiunto un limite superiore oi al numero di transizioni verso l'alto che un dispositivo può subire:

$$\sum_{t=2}^T 1_{S_i}^T u_i(t) \leq o_i \quad \forall i \in \mathcal{N}$$

Tipicamente l'energia elettrica consumata da un dispositivo durante l'orizzonte temporale T è limitata, per modellare questo comportamento si costringe ogni dispositivo i a consumare al più una determinata quantità di potenza:

$$\sum_{t=1}^T p_i^T x_i(t) \leq m_i \quad \forall i \in \mathcal{N}$$

dove il valore m_i rappresenta un limite superiore al consumo specifico per quell'elettrodomestico.

Anziché fare un solo vincolo per dispositivo nel periodo di tempo $[1, T]$, una serie di vincoli più rigidi potrebbero essere applicati partizionando l'orizzonte temporale in un numero fissato di intervalli di tempo. Questo potrebbe comportare una migliore caratterizzazione dei diversi modelli di consumo che si possono avere a seconda della fascia oraria e delle abitudini di coloro che popolano l'edificio. L'intuizione è che gli apparecchi che tipicamente vengono utilizzati durante il giorno difficilmente verranno utilizzati tra l'una e le cinque del mattino; di conseguenza, il seguente vincolo impone che il consumo di energia in un periodo di tempo sia inferiore o uguale ad un consumo massimo m_{ih} consentito per tale periodo:

$$\sum_{t \in \mathcal{T}_h} p_i^T x_i(t) \leq m_{ih} \quad \forall h \in \mathcal{H}, \forall i \in \mathcal{N}$$

dove $\mathcal{H} = \{1, \dots, H\}$ è l'insieme di H intervalli temporali, mentre \mathcal{T}_h è l'insieme di indici $t \in T$ associati al h -esimo intervallo temporale. Per esempio, per i dispositivi asciugatrice e lavastoviglie, la suddivisione appena descritta risulta essere molto utile perché permette di evitare che questi apparecchi, tipicamente usati durante le ore diurne, vengano attivati eccessivamente durante la notte.

Il modello complessivo, ottenuto come spiegato nelle sezioni precedenti, viene riassunto di seguito.

$$\begin{aligned} \min \quad & \sum_{t=1}^T \left(y(t) - \sum_{i=1}^N p_i^T x_i(t) \right)^2 \\ & + \lambda_1 \sum_{t=2}^T \sum_{i=1}^N w_i \|x_i(t) - x_i(t-1)\|_2^2 \\ & + \lambda_2 \sum_{t=1}^T \sum_{i=1}^N l_i (1 - s_i(t)) \|x_i(t)\|_2^2 \\ \text{s.t.} \quad & \mathbf{1}_{S_i}^T x_i(t) \leq 1 \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \\ & \mathbf{1}_{S_i}^T x_i(t) = 1 \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{O} \\ & x_i(t) - x_i(t-1) = \mathbf{u}_i(t) - \mathbf{d}_i(t) \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \\ & \mathbf{u}_i(t) + \mathbf{d}_i(t) \leq \mathbf{1}_{S_i} \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \\ & \mathbf{U}_i \mathbf{u}_i(t) = \mathbf{D}_i \mathbf{d}_i(t) \quad \forall t \in \mathcal{T}, i \in \mathcal{F} \\ & \sum_{\tau=t}^{t+a_i^j-1} x_i^j(\tau) \geq a_i^j w_i^j(t) \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N}, \forall j \in \mathcal{S}_i \\ & \sum_{\tau=t}^{t+b_i^j} x_i^j(\tau) \leq b_i^j \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N}, \forall j \in \mathcal{S}_i \\ & \sum_{t=2}^T \mathbf{1}_{S_i}^T \mathbf{u}_i(t) \leq o_i \quad \forall i \in \mathcal{N} \\ & \sum_{t \in \mathcal{T}_h} p_i^T x_i(t) \leq m_{ih} \quad \forall h \in \mathcal{H}, \forall i \in \mathcal{N} \\ & \sum_{i=1}^N p_i^T x_i(t) \leq y(t) \quad \forall t \in \mathcal{T} \\ & x_i(t), \mathbf{u}_i(t), \mathbf{d}_i(t) \in \{0, 1\}^{S_i} \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \end{aligned}$$

Il modello appena descritto è un problema di ottimizzazione con variabili binarie, funzione obiettivo quadratica e vincoli lineari che quindi rientra nella categoria di problemi nota come Binary Quadratic Programming o BQP; questa è una classe generale di problemi di ottimizzazione notoriamente molto difficili da risolvere a causa della non convessità e dell'interrezza delle variabili. In effetti, questa istanza diventa costosa da risolvere man mano che l'orizzonte temporale T e il numero di dispositivi N aumentano; tuttavia, questa formulazione è risolvibile in modo efficiente per i dati ottenuti da sensori a bassa frequenza in cui la risoluzione varia da 1 a 15 minuti. I dati ottenuti da sensori che lavorano alla risoluzione del minuto sono più pratici da trovare grazie al costo relativamente basso dell'hardware; inoltre avere accesso a questo tipo di dati è più in linea con l'applicazione nel mondo reale del servizio NILM. Per i dati dei contatori intelligenti a bassa frequenza, il nostro problema può essere risolto in ordine di secondi all'ottimale globale da risolutori di programmazione intera all'avanguardia come Gurobi [24].

4.1 Risultati di BQP sui nuovi dati

In questa sezione riportiamo i risultati dell'applicazione di BQP sul dataset costituito dalle case H1, H3, H4, H6, H9 e H11. I dati sono stati aggregati a 5 minuti per mantenere contenuta la complessità computazionale della risoluzione del problema quadratico. L'orizzonte di disaggregazione è stato scelto pari a una settimana. Tre settimane di dati (quelle relative al mese di giugno) sono state utilizzate per la fase di scelta dei parametri, mentre la disaggregazione sul test set di un mese (mese di luglio) è stata effettuata una settimana per volta. I due coefficienti in funzione obiettivo λ_1 e λ_2 sono stati scelti in cross validation. Il problema di ottimizzazione è stato risolto tramite il software Gurobi. Per ogni abitazione e per ogni elettrodomestico riportiamo il valore medio sul periodo di test (1 mese) delle metriche Estimation Accuracy e F1 Score.

H1	Estimation Accuracy	F1 Score
Asciugatrice	0.500	1.000
Computer ragazzi	0.923	0.882
Computer studio	0.905	0.995
Condizionatore matrimoniale	0.901	0.984
Condizionatore soggiorno	0.864	0.930

Condizionatore ragazzi	0.868	0.945
Lavastoviglie	0.914	0.898
Lavatrice	0.697	0.627
Scaldabagno pompa calore	0.897	0.999
TV camera	0.924	0.880
TV soggiorno	0.901	0.857

Tabella 32: risultati di BQP sulla casa H1

H3	Estimation Accuracy	F1 Score
Lavastoviglie	0.852	0.702
TV	0.976	0.988
Lavatrice	0.716	0.915

Tabella 33: risultati di BQP sulla casa H3

H4	Estimation Accuracy	F1 Score
Condizionatore	0.851	0.965
Frigorifero	0.841	0.884
Lavastoviglie	0.892	0.792

Lavatrice	0.750	0.798
------------------	-------	-------

Tabella 34 risultati di BQP sulla casa H4

H9	Estimation Accuracy	F1 Score
Asciugatrice	0.922	0.955
Lavastoviglie	0.847	0.756
Lavatrice	0.794	0.527
Condizionatore camera 1	0.788	0.748
Condizionatore camera 2	0.876	0.938

Tabella 35 risultati di BQP sulla casa H9

HOME 11	Estimation Accuracy	F1 Score
Condizionatore letto	0.922	0.971
Frigorifero	0.922	0.911
Lavastoviglie	0.900	1.000

Tabella 36 risultati di BQP sulla casa H11

Si nota come i risultati siano ottimi, le metriche sia relative all'accuratezza sui consumi che relative al rilevamento delle attivazioni sono molto alte, nonostante la piccolissima quantità di dati utilizzata per il training (solo tre settimane). Per avere una migliore percezione della bontà dei risultati riportiamo nelle Figure successive anche i plot dei pattern predetti per alcuni elettrodomestici particolarmente significativi. Si vede che il metodo funziona benissimo sia per

elettrodomestici always on come il frigo che su elettrodomestici a utilizzo sporadico come lavatrice e lavastoviglie

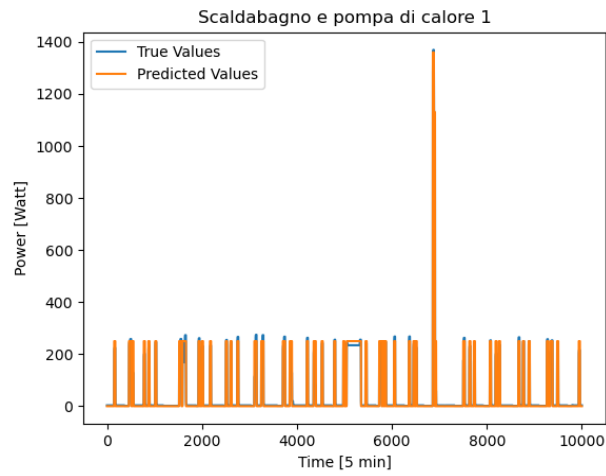


Figura 52: Plot del test set dello scaldabagno di casa H1

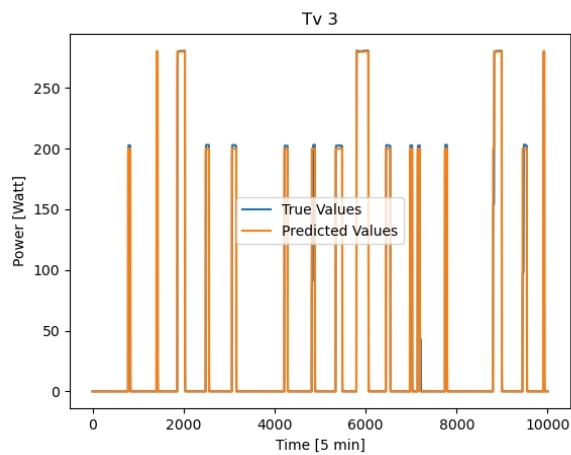


Figura 53: Plot del test set della TV di casa H1

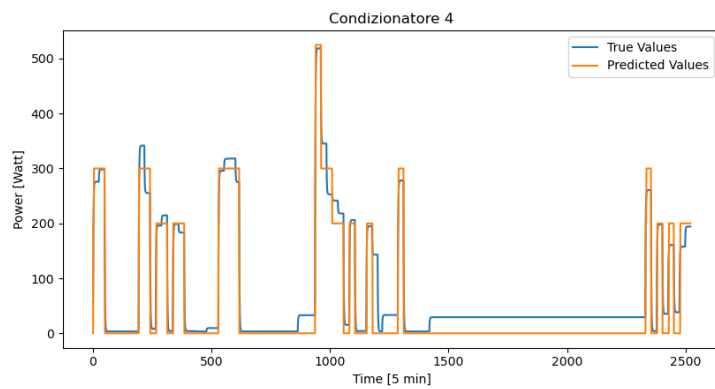


Figura 54: Plot del test set del condizionatore di casa H4

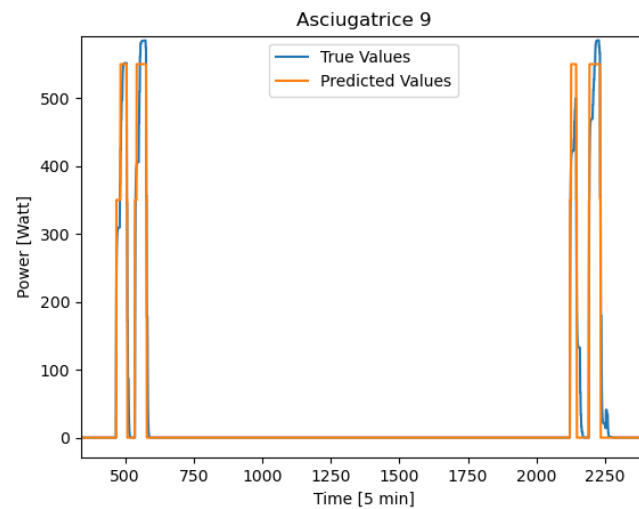


Figura 53: Plot del test set dell'asciugatrice di casa H9

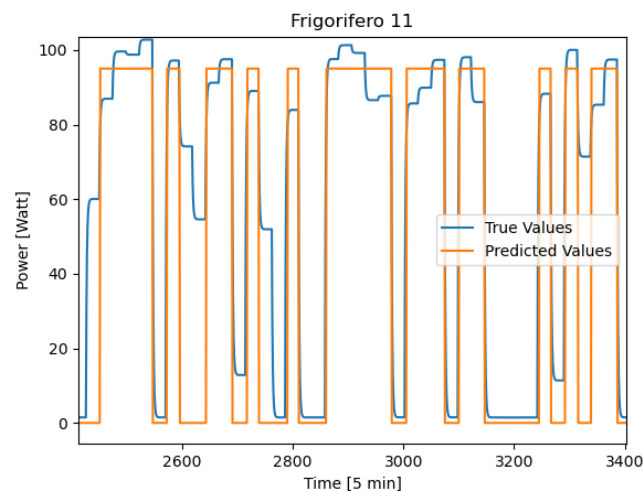


Figura 53: Plot del test set del frigo di casa H11

In conclusione, anche su questo nuovo dataset è stato prodotto una disaggregazione dei consumi ottima a una granularità (5 minuti) che permette di catturare tutti i comportamenti energetici degli abitanti delle diverse abitazioni.

Conclusioni

Durante il progetto sono stati affrontati due temi principali correlati ai consumi energetici: la disaggregazione dei consumi elettrici, e la previsione dei consumi elettrici complessivi. Entrambi i temi possono contribuire a indirizzare i consumatori su comportamenti virtuosi. In particolare, sono stati utilizzati due approcci diversi per la disaggregazione dei consumi corrispondenti a dati diversi disponibili. Quando si ha un insieme di dati molto ampio e a granularità di secondi, la scelta che porta a una maggiore accuratezza è l'utilizzo di una rete deep. Questo implica l'addestramento di una rete deep per ogni elettrodomestico con un insieme di dati di training molto numeroso, con un elevato costo computazionale ma con ottimi risultati in termini di accuratezza. Questo approccio è stato adottato con successo sul primo dataset, contenente i dati di alcune case di Roma. Verso la fine del progetto abbiamo ricevuto un secondo dataset contenente invece un insieme di dati di dimensioni contenute, di case con numerosi elettrodomestici monitorati. Su questo secondo dataset, l'approccio basato su reti deep non è praticabile. Abbiamo quindi utilizzato un approccio

basato sull'ottimizzazione, in cui il problema della disaggregazione viene formulato come un problema quadratico a variabili binarie. Il vantaggio di questo approccio è che richiede un insieme di dati molto piccolo per stimare i parametri del problema di ottimizzazione. I risultati ottenuti hanno dimostrato l'efficacia del metodo.

Per quel che riguarda la previsione dei consumi, si è prima proceduto alle costruzioni di cluster di abitazioni con comportamento energetico simile, in modo da addestrare un unico modello di previsione per ogni profilo di consumo. Le features che si sono dimostrate più rilevanti ai fini dell'individuazione dei profili di consumo sono state features relative ai consumi e utilizzi medi dei diversi elettrodomestici, estratte da alcuni questionari compilati dai residenti, nonché il numero di abitanti e alcune informazioni sulle abitudini di consumo. Una volta costruiti i cluster, le migliori previsioni dei consumi sono state ottenute combinando due predittori dei consumi basati su gradient boosting e su una rete neurale LSTM con un classificatore dei picchi di consumi, capace di correggere la tendenza dei predittori a sottostimare i picchi. I risultati ottenuti dimostrano l'efficacia della fase di clustering in quanto il profilo costruito sulla casa più rappresentativa del cluster funziona bene anche sulle altre case appartenenti allo stesso cluster.

Riferimenti bibliografici

1. G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.
2. M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE transactions on Consumer Electronics*, vol. 57, no. 1, pp. 76–84, 2011.
3. A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.
4. K. He, L. Stankovic, J. Liao, and V. Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1739–1747, 2016.
5. T. Hassan, F. Javed, and N. Arshad, "An empirical investigation of vitrajectory based load signatures for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 870–878, 2013.
6. D. Chowdhury and M. M. Hasan, "Non-intrusive load monitoring using ensemble empirical mode decomposition and random forest classifier," in *Proceedings of the International Conference on Digital Image and Signal Processing (DISP)*, Oxford, UK, 2019, pp. 29–30.
7. Y.-H. Lin and M.-S. Tsai, "Non-intrusive load monitoring by novel neuro-fuzzy classification considering uncertainties," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2376–2384, 2014.
8. G. Elafoudi, L. Stankovic, and V. Stankovic, "Power disaggregation of domestic smart meter readings using dynamic time warping," in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*. IEEE, 2014, pp. 36–39.

9. B. Liu, W. Luan, and Y. Yu, "Dynamic time warping based non-intrusive load transient identification," *Applied energy*, vol. 195, pp. 634–645, 2017.
10. B. Zhao, K. He, L. Stankovic, and V. Stankovic, "Improving eventbased non-intrusive load monitoring using graph signal processing," *IEEE Access*, vol. 6, pp. 53 944–53 959, 2018.
11. S. M. Tabatabaei, S. Dick, and W. Xu, "Toward non-intrusive load monitoring via multi-label classification," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 26–40, 2016.
12. J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Artificial intelligence and statistics*. PMLR, 2012, pp. 1472–1482.
13. W. Kong, Z. Y. Dong, D. J. Hill, J. Ma, J. Zhao, and F. Luo, "A hierarchical hidden markov model framework for home appliance modeling," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3079–3090, 2016.
14. R. Bonfigli, E. Principi, M. Fagiani, M. Severini, S. Squartini, and F. Piazza, "Non-intrusive load monitoring by using active and reactive power in additive factorial hidden markov models," *Applied Energy*, vol. 208, pp. 1590–1607, 2017.
15. O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 26, no. 1, 2012.
16. S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram, "Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2575–2585, 2015.
17. A. Rahimpour, H. Qi, D. Fugate, and T. Kuruganti, "Non-intrusive energy disaggregation using non-negative matrix factorization with sumto-k constraint," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4430–4441, 2017.
18. J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015, pp. 55–64.
19. C. Zhang, M. Zhong, Z. Wang, N. Goddard, and C. Sutton, "Sequenceto-point learning with neural networks for non-intrusive load monitoring," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
20. M. Kaselimi, N. Doulamis, A. Voulodimos, E. Protopapadakis, and A. Doulamis, "Context aware energy disaggregation using adaptive bidirectional lstm models," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3054–3067, 2020.
21. V. Piccialli and A. M. Sudoso, "Improving non-intrusive load disaggregation through an attention-based deep neural network," *Energies*, vol. 14, no. 4, p. 847, 2021.

22. C. Shin, S. Rho, H. Lee, and W. Rhee, "Data requirements for applying machine learning to energy disaggregation," *Energies*, vol. 12, no. 9, p. 1696, 2019.
23. Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso". *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
24. Gurobi Optimization, LLC (2021). "Gurobi Optimizer Reference Manual", "https://www.gurobi.com".
25. L. M. Candanedo, V. Feldheim e D. Deramaix. "Data driven prediction models of energy use of appliances in a low-energy house". In: *Energy Buildings* 140, pp. 81–97, 2017.
26. Richard E. Edwards, Joshua New e Lynne E. Parker. "Predicting future hourly residential electrical consumption: a machine learning case study". *Energy and Buildings* 49, pp. 591–603, 2012.
27. Z.Wang e Y. Ding. "An occupant-based energy consumption prediction model for office equipment». *Energy Buildings* 109, pp. 12–22, 2015.
28. Rishree K. Jain et al. "Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy". *Applied Energy* 123 (2014), pp. 168–178.
29. Katarina Grolinger et al. "Energy forecasting for event venues: Big data and prediction accuracy". *Energy and Buildings* 112, pp. 222–233, 2016.
30. R.K. Jain et al. "Forecasting Energy Consumption of Multi-Family Residential Buildings using Support Vector Regression». *Applied Energy* 123, pp. 168–178, 2014.
31. A. Almalaq e J. Zhang. «Evolutionary deep learning based energy consumption prediction for buildings». *IEEE Access* 7, pp. 1520–1531, 2019.
32. F. Iglesias e W. Kastner. "Analysis of similarity measures in times series clustering for the discovery of building energy patterns». *Energies* 6, pp. 579–597, 2013.
33. J. du Toit et al. "Customer segmentation using unsupervised learning on daily energy load profiles". *Adv Inform Technol* 7, pp. 69–75, 2016.
34. Aven Satre-Meloy, et al. "Cluster analysis and prediction of residential peak demand profiles using occupant activity data». *Applied Energy* 260, 2020.
35. ENEA. KiloWattene-Illuminazione-FAQ. url: <http://kilowattene.enea.it/KiloWattene-lighting-info.html>
36. Lloyd, S. (1982). Least squares quantization in PCM. *IEEE transactions on information theory*, 28(2), 129-137

37. T.K. Wijaya et al. "Cluster-based aggregate forecasting for residential electricity demand using smart meter data". In: Big Data IEEE International Conference (2015), pp. 879–887
38. World Weather Online. url: <https://www.worldweatheronline.com>
39. F. U. M. Ullah et al. "Short-Term Prediction of Residential Power Energy Consumption via CNN and Multi-Layer Bi-Directional LSTM Networks". IEEE Access 8 (2020), pp. 123369–123380.
40. Regressor Chain. url: <https://scikit-learn.org/stable/modules/generated/sklearn.multioutput.RegressorChain.html>
41. F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: Journal of Machine Learning Research 12 (2011), pp. 2825–2830.
42. Jerome H. Friedman. "Greedy Function Approximation: a Gradient Boosting Machine". In: The annals of statistics 29 (2001), pp. 1189–1232.
43. Yoav Freund e Robert E. Schapire. «A Short Introduction to Boosting». Journal of Japanese Society for Artificial Intelligence (1999)
44. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
45. M. Balletti, V. Piccialli, A. Sudoso. "Mixed-Integer Nonlinear Programming for State-based Non-Intrusive Load Monitoring" arXiv:2106.09158, accettato su IEEE Transactions on Smart Grids

Breve curriculum del gruppo scientifico di lavoro:

Gianpaolo Oriolo

Gianpaolo Oriolo è attualmente professore Ordinario di Ricerca Operativa presso l'Università di Roma Tor Vergata, Facoltà d'Ingegneria. Ha conseguito il Dottorato in Ricerca Operativa, Facoltà d'Ingegneria, presso Università di Roma La Sapienza. E' stato postdoc presso London School of Economics, London, UK e presso Center for Discrete and Applicable Mathematics, Bell Laboratories NJ, US. E' Area editor (Graphs and Networks), della rivista Operations Research Letters. Ha vinto l'IBM Faculty award e la borsa di studio DoNet (Discrete Optimization Network). Inoltre ha l'abilitazione alla funzione di professore universitario in Francia. I suoi interessi di ricerca coprono l'ottimizzazione discreta, combinatorica poliedrale, e il network design. E' autore di circa 40 pubblicazioni su riviste internazionali con processo di revisione su riviste quali: Combinatorica, Journal of Graph Theory, Mathematics of Operations Research, Siam Journal on Discrete Mathematics, Mathematical Programming, Discrete Mathematics, Operations Research Letters, Discrete Applied Mathematics, Theoretical Computer Sciences, Networks.

Veronica Piccialli

Veronica Piccialli ha conseguito la laurea in Ingegneria Informatica (summa cum laude) e il dottorato in Ricerca Operativa presso Sapienza Università di Roma rispettivamente nel 2000 e nel 2004. Dal 2008 al 2020 è stata Ricercatore presso l'Università di Roma Tor Vergata. Dal 2020 è stata Professore Associato presso l'Università di Roma Tor Vergata, dove ha insegnato Metodi di Ottimizzazione per Big Data e Machine Learning per la laurea magistrale in Ingegneria Informatica. Dal 2021 è Professore Ordinario presso il Dipartimento di Ingegneria Informatica, Automatica e Gestionale Antonio Ruberti, Sapienza Università di Roma.

Nel 2006 è stata postdoc presso il dipartimento di Combinatorics & Optimization dell'Università di Waterloo, Canada. Nel 2013 ha ottenuto l'abilitazione scientifica nazionale italiana come professore associato. Nel 2017 ha ottenuto l'abilitazione scientifica nazionale come professore ordinario.

Dal 2019 è Associate Editor nell'area "Design and Analysis of Algorithms: Continuo" per INFORMS Journal on Computing. Dal 2021 è Associate Editor per EURO Journal on Computational Optimization. I suoi interessi di ricerca includono l'apprendimento automatico con applicazioni in Brain Computer Interfaces e disaggregazione dei consumi elettrici, teoria dei giochi, programmazione semidefinita, programmazione non lineare mista-intera, ottimizzazione globale con applicazioni in ingegneria chimica e programmazione quadratica standard.

È autrice e co-autrice di più di 40 articoli su riviste rinomate come Mathematical Programming, SIAM Journal on Optimization, IEEE Transactions on Neural Networks and Learning Systems, Computational Optimization and Applications, Computational Management Science, Operations Research Letters, Journal of Global Optimization, Optimization Methods and Software, Journal of Optimization Theory and Applications, ed è autrice di 3 capitoli di libri con referaggio pubblicati in libri internazionali.

Antonio Maria Sudoso

Antonio Maria Sudoso ha conseguito la Laurea Magistrale in Ingegneria Informatica (summa cum laude) nel 2018 presso l'Università degli Studi di Roma Tor Vergata. Dal 2018 al 2021 ha partecipato al Dottorato di Ricerca in Computer Science, Control e Geoinformation ed ha inviato ai revisori una tesi dal titolo 'Global Optimization for Machine Learning Applications'. Dal 2022 è Assegnista di Ricerca presso Sapienza Università di Roma. I suoi interessi di ricerca si collocano all'intersezione tra il machine learning e l'ottimizzazione.