Data mining Coursework 2023

Task: 1 Write a literature survey of the state of the art, using up to 5000 words, excluding figures, tables, and references for one of the topics below. Extensive use of recent bibliography and review of several recent research papers is strongly encouraged.
Topic chosen: Data mining for Energy Forecasting

1 Introduction

In today's rapidly evolving world, where energy demand continues to rise, and resources become scarcer, accurate forecasting of energy needs has become crucial for effective decision-making and sustainable resource management. To meet these challenges, the application of data mining techniques has emerged as a vital tool in the field of energy forecasting. Data mining, a powerful analytical process that extracts valuable knowledge and patterns from vast amounts of data, offers numerous benefits and plays a significant role in ensuring efficient energy planning, optimal resource allocation, and the overall advancement of the energy sector. We will explore why it is of paramount importance to employ data mining for energy forecasting and delve into the key reasons behind its indispensability in shaping a more sustainable and resilient energy future.

Fossil fuels are still the most important energy source in the world today. Fossil fuels are hydrocarbons or their derivatives, including natural resources such as coal, oil and natural gas. Fossil fuels take millions of years to form, and known usable reserves are depleting much faster than new fossil fuels are produced. At the same time, fossil fuels emit greenhouse gases that accelerate climate change, including global warming, and threaten the environment we depend on. Renewable energy has therefore received a great deal of attention around the world in recent years. Renewable energy refers to energy that can be recycled in nature, such as solar, wind, hydropower, and geothermal. Renewable energy has at least two advantages over fossil fuels. First, the world's renewable energy resources are abundant, renewable and inexhaustible. Second, renewable energy is clean, green, low-carbon and helps protect the environment. Specifically, renewable energy effectively reduces sulfide (SO2), carbide (CO) and dust emissions, reducing the risk of air pollution and the greenhouse effect. In addition, the use of renewable energy can achieve the objective of reducing the use of natural fossil fuels and protecting the ecological environment. In addition, renewable energy reduces solid waste emissions and reduces soil pollution. [1] Renewable energy can also reduce exhaust gas and waste liquid emissions during use, so it can also achieve the goal of protecting water resources. Renewable energy has therefore developed very rapidly in recent years [2]. According to the REN21 2017 report, renewable energy accounted for 19.3% of global energy consumption and 24.5% of electricity generation in 2016[3]. Many countries, such as the United States and China, have developed various regulations, incentives, and subsidies to promote renewable energy deployment [4].

The study of energy consumption issues has become an important research topic in recent decades. According to economic theory, energy is one of the most important resources for industrial production, and forecasting energy consumption is an important step in macro-planning for industry and the energy sector [5]. The long-term planning of energy demand and supply must meet the country's sustainable development requirements. Accurate forecasts help decision makers understand future energy consumption and trends to better plan and plan public system operations. In recent decades, population growth and industrial economic development have increased energy consumption in many underdeveloped countries. Today, energy has become an important factor for the economic and even socio-economic development of each country [6].

Efficient planning of energy allocation requires an accurate forecast of future demand to balance energy supply and demand [7]. Therefore, using models to accurately predict future energy consumption trends, especially using nonlinear data, is an important issue for power generation and distribution systems [8]. Basic operations of the energy system, such as economic scheduling, unit maintenance, fuel planning, and unit deployment, can be performed more efficiently with more accurate forecasts [9]. Predictive research is also important from a price perspective. Fluctuations in energy prices are determined by the balance between supply and demand. Fluctuations in energy prices, climate change, increasing global energy demand, dependence on fossil fuels, and declining new energy development threaten energy security [10]. Environmental issues such as global warming and emissions are also important aspects of energy consumption forecasting.

2 State of the art

This section reports state-of-the-art Machine Learning (ML) and Deep Learning (DL) algorithms for energy forecasting.

2.1 Machine learning

Machine learning techniques have proven to be invaluable in the domain of energy forecasting, enabling accurate and reliable predictions in the face of complex and dynamic energy systems. By analyzing vast amounts of historical energy data, machine learning algorithms can uncover hidden patterns, correlations, and trends that may not be readily apparent to human analysts. These algorithms are capable of automatically learning from the data, adapting to changing conditions, and continuously improving their forecasting accuracy over time. With the ability to handle diverse data sources such as weather patterns, electricity demand, renewable energy generation, and market dynamics, machine learning models can capture the intricate interdependencies between these factors and provide insightful predictions for energy demand, supply, and price fluctuations. This empowers energy stakeholders to make informed decisions, optimize resource allocation, improve energy efficiency, and plan for a more sustainable and resilient future.

2.1.1 Support Vector Regression

Support Vector Regression (SVR): SVR is a supervised learning algorithm, used mainly for classification tasks, that has been widely used for energy forecasting. It can effectively capture complex relationships between historical energy data and other relevant variables.

A data-driven approach enables highly accurate predictions of energy consumption. Among them, his SVR, which applies the support vector machine (SVM) proposed by Vapnik [11] to regression analysis, has a wide range of applications in energy prediction field. SVR is based on the principle of structural risk minimization. It has its own advantages for small datasets and can maintain good generalization ability. However, according to the literature, SVR-based models for predicting building energy consumption are mainly characterized by two factors:

SVR hyperparameters and kernel functions. Regarding the structure of the SVR, the hyperparameters determine the support vectors and the kernel functions determine the properties of the high-dimensional feature space.

The Ref. [12] to ensure the reliability of data-driven models for predicting building energy consumption, their accuracy, generalizability and robustness need to be improved. A new vector field-based support vector regression (SVR) is proposed to overcome the above limitations. By warping the sample data space or high-dimensional feature space using a vector field, this technique transforms the highly nonlinearity between input and output into linearity, ensuring robustness and generalizability of the model can improve the prediction accuracy. This study shows that improving algorithms can further improve the prediction accuracy of data-driven models for predicting building energy consumption. By applying multi-distortion to either the sample space or feature space, the inherent nonlinearity present in the dataset is transformed into linearity. This approach tackles the challenge of constructing a highly complex nonlinear model and offers a theoretical solution to overcome this limitation.

The Ref. [13] discussed the utilization of hybrid intelligent computing methods and swarm-based algorithms, in conjunction with the support vector regression (SVR) model, shows significant promise in addressing the issue of premature convergence. In this study, a new electric load forecasting model based on SVR is proposed. This model combines variational mode decomposition (VMD), the chaotic mapping mechanism, and the grey wolf optimizer (GWO) within the VMD-SVR-CGWO framework. The objective is to enhance the search process and find optimal combinations of SVR parameters that lead to improved forecasting accuracy. Experiments conducted on two well-known electric load datasets demonstrate that the VMD-SVR-CGWO model, as proposed in this study, exhibits superior performance in terms of forecasting accuracy when compared to other models.

The Ref. [14] about wind speed forecasting used SVR known for its strong generalization capabilities, to predict the detailed components of the dataset.

2.1.2 Random Forest

Random Forest (RF): Random Forest is an ensemble learning method that combines multiple decision trees. It is commonly applied to energy forecasting tasks, leveraging features such as historical energy consumption, weather data, and other factors.

The Ref. [15] author of the study addressed the increasing share of photovoltaic power generation in the power trading market, which poses challenges due to its intermittent and uncontrollable nature, leading to potential instability in the power system. To enhance the prediction accuracy of photovoltaic power generation, the author proposed a novel approach that combines Principal Component Analysis (PCA) and K-means clustering with a random forest algorithm optimized using the Differential Evolution Grey Wolf Optimizer. This methodology aims to mitigate issues such as excessive noise in the original data and improper parameter adjustment commonly observed in traditional modeling processes. By applying PCA and K-means clustering, hourly point features like the predicted time points are obtained, facilitating the filtering of input data to reduce interference from noisy data. Moreover, leveraging the efficiency of the optimization algorithm, the proposed model effectively selects the parameters of the random forest algorithm, reducing the impact of artificial filtering factors and associated errors. Comparative experiments conducted in three regions demonstrate that the recommended model exhibits higher prediction accuracy and robustness, indicating its potential for improving photovoltaic power generation forecasting.

Ref. [16] discussed about accurately forecasting the impact of different usage patterns across various load types on energy consumption is crucial for reducing carbon emissions and implementing effective demand-side energy management strategies. The study focuses on examining the medium-term (MT) and long-term (LT) energy predictions for utilities, independent power producers, and industrial customers, aiming to estimate the energy usage requirements for large-scale city-wide scenarios. The nonlinear autoregressive model (NARM), linear model using stepwise regression (LMSR), and random forest (LSBoost) techniques are employed, utilizing real-world environmental and energy consumption data. To address irregular load patterns and eliminate abnormal trends in energy usage, outlier detection and clustering analysis techniques are applied. This enables the identification and removal of anomalies, facilitating more accurate energy forecasting and enabling a better understanding of the impact of climate changes on energy consumption.

The Ref. [17] found that maintaining a balanced carbon cycle is crucial for ensuring the sustainability and health of ecosystems. Net ecosystem carbon exchange (NEE), which reflects the carbon balance between biological organisms and the atmosphere, is influenced by various meteorological variables to varying degrees. This study focuses on predicting NEE using data collected from two flux measuring sites. The gradient boosting regression algorithm is employed, utilizing meteorological and flux data from the UK-Gri site. To prevent overfitting, KFold cross-validation is implemented during the training process, while the random forest algorithm is utilized to identify the key variables that predominantly influence NEE. Among these variables, global radiation, photosynthetic active radiation, minimum soil temperature, and latent heat are identified as the most important. The performance of the regression model is compared to three state-of-the-art prediction models: support vector machine, stochastic gradient descent, and Bayesian ridge. Experimental results demonstrate that the proposed regression model outperforms the other models, exhibiting higher R-squared values, lower mean absolute error, and root mean squared error. To assess the model's generalization ability, data from a second flux site (NL-Loo) is employed, and the hybrid data from both sites is utilized. The results indicate that the model performs well with the hybrid data as well. In practical terms, the gradient boosting regression model offers tunable hyperparameters and loss functions, making it more flexible and accurate compared to the alternative models. This study establishes the value of combining gradient boosting regression with random forest models for effective NEE prediction and identification of influential variables. The methodologies presented have potential applications in ecosystem stability evaluation, environmental restoration, climate change trend analysis, and global warming monitoring.

2.1.3 Gradient Boosting Machines

Gradient Boosting Machines (GBM): GBM is an ensemble learning technique that combines weak predictive models, typically decision trees, to create a strong predictive model. GBM has been successfully utilized in energy forecasting to improve accuracy and handle complex relationships between variables, also XGBoost (eXtreme Gradient Boosting) is an optimized implementation of gradient boosting machines (GBM) and has gained popularity in various domains, including energy forecasting. XGBoost is known for its ability to handle large datasets, complex relationships, and feature interactions. It combines multiple weak predictive models, typically decision trees, in an iterative manner, aiming to create a strong ensemble model that can accurately predict energy demand or supply.

The Ref. [18] developed a new model for predicting the electricity consumption of buildings, with the aim of improving accuracy. This innovative approach uses the Gradient Boosting regression tree algorithm to simulate and predict energy consumption. Experimental results comparing different forecasting models reveal that the proposed model achieves lower root mean square error (RMSE) and mean absolute error (MAE) values on different test datasets. Moreover, extensive experiments conducted to compare the proposed model with existing alternatives demonstrate its superiority in accurately predicting energy consumption.

The Ref. [19] proposed a novel hybrid model which is a combination of CEEMDAN and XGBoost. The raw data is divided into several refined datasets using a technique called complete ensemble empirical mode decomposition with adaptive noise. Subsequently, the building energy consumption is forecasted using the traditional extreme gradient boosting method. The simulation focuses on the daily energy consumption of the City of Bloomington Intake Tower. The outcomes indicate that the proposed model achieves a mean absolute percentage error of 4.85%, which is significantly lower than the errors obtained by five benchmark models, see Fig.1 below. Furthermore, the proposed model demonstrates favorable predictive capabilities when applied to other parameters associated with the intake tower's energy consumption.

Model	MAE (kWh)	RMSE (kWh)	MAPE (%)	RMSPE (%)	U1	U2
CEEMDAN-XGBoost	717.39	938.43	4.85	6.46	0.030	0.061
XGBoost	1227.15	1518.24	8.06	10.15	0.049	0.099
CEEMDAN-RF	941.92	1278.08	6.26	8.59	0.041	0.083
RBFNN	1176.83	1472.99	7.67	9.69	0.048	0.096
PSO-SVM	1209.18	1508.61	7.92	10.05	0.049	0.098
LSSVM	1199.88	1485.41	7.87	9.86	0.048	0.096

Fig 1: Prediction errors of CBU Intake Tower daily energy consumptions [19]

The Ref. [20] developed in this study, a novel ensemble model named XGBF-DNN for predicting hourly global horizontal irradiance. The ensemble model combines two advanced base models, namely extreme gradient boosting forest and deep neural networks, using ridge regression to address the issue of over-fitting. The framework is carefully designed to ensure diversity among the base models, which is recognized as crucial for the success of an ensemble model. To enhance accuracy, a subset of relevant input features including temperature, clear-sky index, relative humidity, and hour of the day is selected. To validate the proposed model comprehensively and reliably, data from three different climatic locations in India are used. Additionally, a seasonal analysis is conducted to gain deeper insights into the model's performance. To evaluate its effectiveness, the proposed model is compared with benchmark methods like smart persistence and traditional machine learning techniques such as random forest, support vector regression, extreme gradient boosting forest, and deep neural networks. The results, see Fig.2 below, demonstrate that the proposed ensemble model, XGBF-DNN, achieves the best combination of stability and prediction accuracy across various seasonal weather conditions. The forecast skill score falls within a range of approximately 33% to 40% in terms of prediction error. These findings highlight the model's predictive performance and stability, making it an ideal and reliable choice for hourly global horizontal irradiance prediction. Consequently, the developed model holds potential for applications in diverse domains, including solar power forecasting, wind power forecasting, electricity consumption prediction, and more.

	MBE	RMSE (W/m²)					FS
	Annual	Annual	Winter	Summer	Monsoon	Autumn	Annual
Jaipur							
SP	−0.363	83.42	102.93	57.32	118.77	52.74	0
SVR	0.705	68.67	81.24	51.81	96.65	45.12	0.176
RF	1.187	67.56	81.68	47.72	94.03	43.68	0.190
XGBoost	0.481	63.09	79.26	46.33	**85.31**	40.53	0.244
DNN	−0.681	**60.85**	**72.18**	**44.64**	90.81	**38.15**	**0.271**
XGBF-DNN (excluding feature selection)	−0.749	56.68	67.92	40.38	78.27	36.32	0.321
XGBF-DNN (including feature selection)	−0.384	53.79	65.81	37.19	72.46	35.47	0.355
New Delhi							
SP	1.255	85.86	103.38	64.39	120.57	60.06	0
SVR	0.916	68.48	86.71	60.07	93.82	53.63	0.204
RF	−1.225	66.01	82.31	56.16	90.03	51.82	0.231
XGBoost	1.392	60.93	80.47	50.73	**83.39**	48.79	0.290
DNN	−0.864	**58.08**	**77.18**	**48.19**	86.16	**48.31**	**0.323**
XGBF-DNN (excluding feature selection)	0.367	53.78	68.43	44.14	81.26	40.21	0.374
XGBF-DNN (including feature selection)	0.0831	51.35	64.14	42.76	77.09	38.61	0.402
Gangtok							
SP	5.895	134.24	97.16	168.48	188.37	76.24	0
SVR	2.109	121.17	90.81	148.64	167.91	70.36	0.097
RF	2.317	118.39	88.01	145.28	160.01	66.15	0.118
XGBoost	−1.935	**105.71**	**82.41**	**131.26**	**149.25**	**61.81**	**0.213**
DNN	1.596	109.34	85.31	138.28	153.18	64.73	0.185
XGBF-DNN (excluding feature selection)	1.107	91.86	80.16	121.25	132.07	58.61	0.316
XGBF-DNN (including feature selection)	0.831	89.13	78.25	119.41	129.06	55.16	0.336

Fig 2: The performance comparison of different studied models on testing dataset [20]

3.1 Deep learning

Deep learning, a subfield of machine learning, has emerged as a powerful tool for energy forecasting due to its ability to automatically extract intricate features and patterns from large-scale and complex energy datasets. Deep learning models, such as deep neural networks, are characterized by their hierarchical architecture consisting of multiple layers of interconnected nodes, known as neurons. These models excel at learning abstract representations and capturing nonlinear relationships within the data, enabling them to handle the intricate dynamics and uncertainties present in energy systems. Deep learning algorithms can process diverse data sources, including historical energy consumption, weather conditions, renewable energy generation, and market variables, to uncover complex dependencies and make accurate predictions. By leveraging the vast computational power of deep learning, energy forecasters can achieve enhanced accuracy, especially in scenarios where traditional forecasting methods may fall short. Deep learning also has the potential to continually improve and adapt its forecasting capabilities as new data becomes available, making it a valuable tool for decision-making, resource planning, and optimizing energy operations.

3.1.1 Long Short-Term Memory

Long Short-Term Memory (LSTM): LSTM is a type of recurrent neural network (RNN) that is effective in capturing long-term dependencies in time series data. It has been widely used for energy demand and supply forecasting due to its ability to model complex patterns.

This article, Ref. [21], introduces kCNN-LSTM, a deep learning framework designed to accurately predict building energy consumption using recorded data at predefined intervals. The kCNN-LSTM approach combines three key components: i) k-means clustering: Utilized for cluster analysis, allowing for the identification of energy consumption patterns and trends. ii) Convolutional Neural Networks (CNN): Employed to extract complex features, capturing non-linear interactions that influence energy consumption. iii) Long Short-Term Memory (LSTM) neural networks: Incorporated to model temporal information and handle long-term dependencies within the time series data.

To evaluate the effectiveness and practicality of kCNN-LSTM, real-time building energy consumption data from a four-story building in IIT-Bombay, India, was utilized. The performance of kCNN-LSTM was compared with the k-means variant of state-of-the-art energy demand forecast models, using well-established quality metrics, see figures below. The results demonstrated that kCNN-LSTM provides accurate energy demand forecasts by effectively capturing spatio-temporal dependencies within the energy consumption data. This establishes kCNN-LSTM as a suitable deep learning model for energy consumption forecasting tasks.

Type	Metrics	Models					
		ARIMA	DBN	MLP	LSTM	CNN	CNN-LSTM
Overall	**MSE**	0.0102	0.0183	0.0189	0.6226	0.3313	**0.0095**
	RMSE	0.0982	0.1354	0.1376	0.7890	0.5756	**0.0974**
	MAE	0.1098	0.1063	0.0998	0.7310	0.5429	**0.0711**
	MAPE	0.8128	0.3550	0.3161	2.0425	1.7206	**0.2697**
Weekdays	**MSE**	0.0356	0.0966	0.0805	0.0267	0.0778	**0.0168**
	RMSE	0.1887	0.3107	0.2837	0.1633	0.2790	**0.1297**
	MAE	0.1527	0.2842	0.2547	0.1240	0.2594	**0.1113**
	MAPE	0.2701	0.8009	1.9979	0.4433	0.7682	**0.3872**
Weekend	**MSE**	0.0293	0.0040	0.0045	0.0048	0.0048	**0.0034**
	RMSE	0.1712	0.0635	0.0667	0.0693	0.0693	**0.0580**
	MAE	0.1471	0.0547	0.0537	0.0570	0.0602	**0.0481**
	MAPE	0.1447	0.1737	0.1580	0.1725	0.1826	**0.1425**
Monday	**MSE**	0.0151	0.0286	0.0073	0.0091	0.0030	**0.0013**
	RMSE	0.1232	0.1690	0.0853	0.0953	0.0551	**0.0364**
	MAE	0.1031	0.1585	0.0580	0.0805	0.0404	**0.026**
	MAPE	0.3119	0.5046	0.1320	0.2486	0.1166	**0.0728**
Wednesday	**MSE**	0.0116	0.0039	0.0023	0.0066	0.0021	**0.00100**
	RMSE	0.1077	0.0624	0.0477	0.0812	0.0455	**0.0324**
	MAE	0.0925	0.0509	0.0398	0.0635	0.0387	**0.0267**
	MAPE	0.2506	0.1635	0.1102	0.1851	0.1128	**0.0765**
Friday	**MSE**	0.3271	0.0180	0.0095	0.0038	0.0042	**0.0024**
	RMSE	0.5720	0.1343	0.0975	0.0618	0.0650	**0.0491**
	MAE	0.5024	0.1229	0.0825	0.0511	0.0551	**0.0368**
	MAPE	0.2444	0.3860	0.1860	0.1455	0.1601	**0.1033**
Sunday	**MSE**	0.0064	0.0022	0.0011	0.0010	0.0009	**0.0008**
	RMSE	0.0803	0.0474	0.0336	0.0324	0.0306	**0.0285**
	MAE	0.0679	0.0419	0.0273	0.0273	0.0252	**0.0232**
	MAPE	0.0966	0.1502	0.0867	0.0924	0.0856	**0.0783**

Fig 3: Cluster 1-Performance analysis of the state-of-the-art machine learning and deep learning energy demand forecast models *[21]*

Type	Metrics	Models					
		ARIMA	DBN	MLP	LSTM	CNN	CNN-LSTM
Overall	**MSE**	0.1006	0.1303	0.1116	0.0618	0.0639	**0.0212**
	RMSE	0.1625	0.3609	0.3341	0.2485	0.2528	**0.1456**
	MAE	0.1138	0.3215	0.2931	0.2133	0.2278	**0.0997**
	MAPE	0.2062	0.2549	0.2375	0.5785	0.2358	**0.2054**
Weekdays	**MSE**	0.1007	0.1007	0.0384	0.0504	0.0687	**0.0251**
	RMSE	0.2066	0.3174	0.1958	0.2246	0.2621	**0.1586**
	MAE	0.1456	0.2718	0.1438	0.1911	0.2206	**0.1397**
	MAPE	0.1918	0.3015	0.4496	0.1855	0.1926	**0.1653**
Weekend	**MSE**	0.0101	0.0416	0.0156	0.0057	0.0334	**0.0038**
	RMSE	0.2134	0.2039	0.1251	0.0757	0.1828	**0.062**
	MAE	0.1329	0.1892	0.1052	0.0579	0.1683	**0.0433**
	MAPE	0.8885	0.5277	0.2774	0.1704	0.4541	**0.1425**
Monday	**MSE**	0.0113	0.0272	0.0230	0.0122	0.0147	**0.0110**
	RMSE	0.1190	0.1648	0.1516	0.1106	0.1211	**0.1050**
	MAE	0.0829	0.1391	0.1230	0.0771	0.0943	**0.0721**
	MAPE	0.2252	0.4564	0.2893	0.2287	0.2838	**0.2169**
Wednesday	**MSE**	0.0094	0.0229	0.0095	0.0099	0.0126	**0.0002**
	RMSE	0.0971	0.1512	0.0975	0.0995	0.1124	**0.0164**
	MAE	0.1040	0.1125	0.0714	0.0708	0.0865	**0.0707**
	MAPE	0.2576	0.4807	0.1733	0.1817	0.2370	**0.1551**
Friday	**MSE**	0.0218	0.1466	0.1293	0.1218	0.1222	**0.0013**
	RMSE	0.1372	0.1478	0.0969	0.1118	0.0967	**0.0871**
	MAE	0.1162	1.0291	0.2317	0.8481	0.9251	**0.1137**
	MAPE	0.2259	0.3521	0.2725	0.2838	0.2915	**0.2346**
Sunday	**MSE**	0.9105	0.0653	0.0167	0.0148	0.0149	**0.0068**
	RMSE	0.0203	0.0048	0.0041	0.0039	0.0046	**0.0036**
	MAE	0.1182	0.0695	0.0640	0.0626	0.0679	**0.0600**
	MAPE	0.3103	0.0542	0.0470	0.0448	0.0512	**0.0418**

Fig 4: Cluster 2- Performance analysis of the state-of-the-art machine learning and deep learning energy demand forecast models [21]

Type	Metrics	Models					
		ARIMA	DBN	MLP	LSTM	CNN	CNN-LSTM
Overall	**MSE**	0.0019	0.0708	0.0380	0.0113	0.0124	**0.0010**
	RMSE	0.1047	0.2660	0.1950	0.1049	0.1114	**0.0303**
	MAE	0.0812	0.2079	0.1450	0.0880	0.0893	**0.0165**
	MAPE	0.2086	0.3882	0.3107	0.1998	0.2101	**0.1670**
Weekdays	**MSE**	0.1006	0.0552	0.4857	0.2612	0.1323	**0.0142**
	RMSE	0.3264	0.2350	0.2204	0.1616	0.1193	**0.1150**
	MAE	0.2180	0.1892	0.1678	0.1348	0.9167	**0.0922**
	MAPE	3.3808	0.3675	0.3686	0.3208	0.2199	**0.2134**
Weekend	**MSE**	0.0024	0.0019	0.0015	0.0016	0.0013	**0.0013**
	RMSE	0.1207	0.0433	0.0389	0.0397	0.0366	**0.0362**
	MAE	0.1138	0.0357	0.0308	0.0304	0.0283	**0.0275**
	MAPE	0.8961	0.0847	0.0702	0.0700	0.0677	**0.0646**
Monday	**MSE**	0.1006	0.0482	0.0417	0.0290	0.0292	**0.0256**
	RMSE	0.3250	0.2196	0.2042	0.1702	0.1708	**0.1601**
	MAE	0.2177	0.1641	0.1566	0.1347	0.1379	**0.1333**
	MAPE	0.2332	0.2704	0.4117	0.2303	0.2337	**0.2257**
Wednesday	**MSE**	0.1006	0.0662	0.0463	0.0314	0.0218	**0.0153**
	RMSE	0.1253	0.2572	0.2151	0.1771	0.1478	**0.1235**
	MAE	0.2169	0.2010	0.1623	0.1443	0.1186	**0.1045**
	MAPE	0.5244	0.3756	0.4868	0.2569	0.2460	**0.1947**
Friday	**MSE**	0.1317	0.0507	0.0578	0.0280	0.0192	**0.0101**
	RMSE	0.3630	0.2251	0.2403	0.1673	0.1386	**0.1005**
	MAE	0.3126	0.1769	0.1885	0.1339	0.1180	**0.0778**
	MAPE	0.4956	0.3336	0.6999	0.2457	0.2384	**0.1473**
Sunday	**MSE**	0.0251	0.0351	0.0201	0.0109	0.0104	**0.0075**
	RMSE	0.1586	0.1873	0.1419	0.1045	0.1018	**0.0867**
	MAE	0.1397	0.1604	0.1144	0.0835	0.0796	**0.0611**
	MAPE	0.3280	0.3643	0.3987	0.1942	0.1929	**0.1404**

Fig 5: Cluster 3- Performance analysis of the state-of-the art machine learning and deep learning energy demand forecast models [21]

The Ref. [22] said that the reliability, stability, efficiency, and accuracy of forecasting methods present significant challenges in the field of energy consumption prediction. Existing forecasting models often struggle with the volatility of energy consumption data, requiring AI models that can effectively predict sudden irregular changes and capture long-term dependencies within the data. To address this, a novel hybrid AI-based forecasting model is introduced in this study, see Fig. 6. The proposed model combines two techniques: singular spectrum analysis (SSA) and parallel long short-term memory (PLSTM) neural networks. The use of SSA for decomposition enhances the performance of the PLSTM network. Through experimental evaluation, the results demonstrate that the proposed model surpasses state-of-the-art models in terms of both prediction accuracy and computational efficiency across various time intervals. In summary, the study presents a hybrid AI-powered forecasting model that effectively addresses the challenges of energy consumption prediction. By combining SSA and PLSTM, the model achieves improved performance compared to existing methods, showcasing enhanced prediction accuracy and computational efficiency.

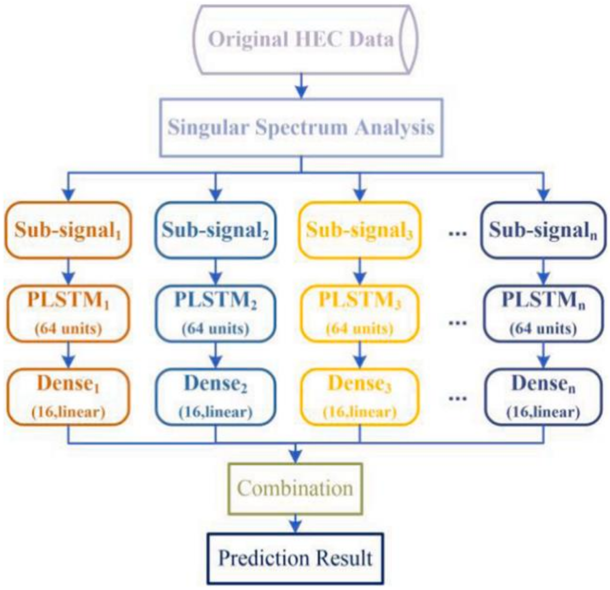

Fig 6: The flow chart of the proposed model [22]

Energy consumption data in real systems often exhibits periodicity, which is not adequately addressed by general forecasting methods. This paper, the Ref. [23] presents a novel approach that utilizes long short-term memory (LSTM) networks to predict periodic energy consumption. The method involves extracting hidden features by analyzing the autocorrelation graph of real industrial data. Correlation analysis and mechanism analysis are employed to identify suitable secondary variables as input for the model. Additionally, the time variable is incorporated to capture the periodic nature more accurately. A LSTM network is then constructed to model and forecast sequential data. Experimental results conducted on a specific cooling system demonstrate that the proposed method outperforms traditional forecasting methods, including autoregressive moving average (ARMA), autoregressive fractional integrated moving average (ARFIMA), and backpropagation neural network (BPNN). The root mean square error (RMSE) of the LSTM model is significantly lower, with reductions of 19.7%, 54.85%, and 64.59% compared to BPNN, ARMA, and ARFIMA, respectively, based on May test data. Furthermore, considering the potential limitations of missing measuring equipment, new prediction models are trained with reduced secondary variables to explore the relationship between prediction accuracy and input variables. The experimental results highlight the excellent generalization capability of the proposed algorithm. In summary, this study introduces an innovative approach that leverages LSTM networks to predict periodic energy consumption. By extracting hidden features, considering appropriate secondary variables, and incorporating the time variable, the proposed method demonstrates superior prediction performance compared to traditional forecasting models. The algorithm's generalization capability is also validated, showcasing its potential for real-world applications.

3.1.2 Convolutional Neural Network

Convolutional Neural Networks (CNN): CNNs, primarily known for their applications in image processing, can also be used for energy forecasting tasks. They can extract useful features from multidimensional data, such as weather patterns and energy consumption data.

The Ref. [24] researchers are drawn to the field of electric energy forecasting due to its significance in energy conservation. Existing mainstream models, such as Gradient Boosting Regression (GBR), Artificial Neural Networks (ANNs), Extreme Learning Machine (ELM), and Support Vector Machine (SVM), encounter challenges in handling the non-linear relationship between input data and output predictions. These models also exhibit limited adaptability in real-world scenarios. Moreover, the energy forecasting domain requires greater robustness, higher prediction accuracy, and improved generalization ability for practical implementation. To address these challenges, this paper proposes a hybrid sequential learning-based energy forecasting model that combines Convolutional Neural Network (CNN) and Gated Recurrent Units (GRU) in a unified framework to achieve accurate energy consumption prediction. The proposed framework consists of two main phases: data refinement and training. In the data refinement phase, preprocessing strategies are applied to enhance the raw data. In the training phase, CNN features are extracted from the input dataset and fed into GRU, which is selected as the optimal choice based on extensive experiments and observed to possess enhanced sequence learning capabilities. The proposed model offers an effective alternative to previous hybrid models in terms of computational complexity and prediction accuracy. This is attributed to the representative feature extraction capabilities of CNNs and the effective gated structure of multi-layered GRU. Experimental evaluation conducted on existing energy forecasting datasets demonstrates the superior performance of the proposed method in terms of precision and efficiency. Notably, the proposed method achieves the smallest error rate on the Appliances Energy Prediction (AEP) dataset, see Fig.7. In summary, this study introduces a hybrid sequential learning-based energy forecasting model that combines CNN and GRU. The model surpasses existing approaches in terms of computational complexity and prediction accuracy, offering improved precision and efficiency. The experimental evaluation validates the superiority of the proposed method, highlighting its potential for accurate energy consumption prediction in real-world applications.

Method	MSE	RMSE	MAE
Linear regression	0.16	0.41	0.30
Decision tree	0.17	0.41	0.33
SVR	0.12	0.35	0.27
CNN	0.17	0.41	0.32
LSTM	0.25	0.50	0.36
CNN-LSTM	0.14	0.38	0.30
Proposed	**0.09**	**0.31**	**0.24**

Fig 7: Performance of different machine learning and deep learning models for AEP dataset [24]

In recent years, the integration of diverse energy sources into power systems has led to the emergence of regional integrated energy systems (IES). However, the coupling and complementation of multiple energy sources present challenges in load forecasting. Given the time-sequence and non-linear characteristics of electric load and the complementarity of different energy sources in IES, this study proposes a hybrid model combining attention-based Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Bidirectional Long Short-Term Memory (BiLSTM) for short-term load forecasting in IES. The model utilizes historical load data, temperature, cooling load, and gas consumption from the previous 5 days as input features. A CNN with an attention block is employed to extract effective features related to load impact factors. Subsequently, the LSTM combined with BiLSTM layers is utilized to forecast the load for the next hour. The proposed model is validated using data from an integrated energy park in North China. The results demonstrate that the proposed method outperforms other models, including CNN-BiLSTM, CNN-LSTM, BiLSTM, LSTM, backpropagation neural network (BPNN), random forest regression (RFR), and support vector machine regression (SVR), in terms of forecasting performance. In summary, this study introduces a hybrid model that combines attention-based CNN, LSTM, and BiLSTM for short-term load forecasting in IES. By leveraging the historical data and employing a multi-layered approach, the proposed model achieves better forecasting performance compared to other models. The validation using data from an integrated energy park supports the effectiveness of the proposed method in load forecasting within the context of IES. [25]

In today's technological era, the demand for electricity is rapidly increasing due to factors such as population growth, electric vehicles, and household appliances. Accurate prediction of energy consumption (ECP) and effective integration of local energy systems (ILES) are crucial for enhancing clean energy management between consumers and suppliers. However, existing approaches for long- and short-term ECP face challenges in dealing with environmental factors and occupant behavior. To address these challenges, we propose a novel hybrid network model called 'DB-Net', which combines a dilated convolutional neural network (DCNN) with bidirectional long short-term memory (BiLSTM). This approach enables efficient control of power energy in ILES between consumers and suppliers for both long- and short-term ECP. The model consists of three phases. In the first phase, a preprocessing module is designed to optimize the collected data and handle outliers through data acquisition and refinement procedures. The goal is to ensure high-quality input data for subsequent processing. In the second phase, the refined data is passed through DCNN layers for feature encoding. The DCNN is capable of capturing important patterns and relationships within the data. The encoded features are then fed into BiLSTM layers, which are responsible for learning hidden sequential patterns and decoding the feature maps. In the final phase, the DB-Net model predicts multi-step power consumption (PC) including hourly, daily, weekly, and monthly outputs. By leveraging the combined power of DCNN and BiLSTM, the model achieves superior predictive performance compared to existing methods, validating its effectiveness in accurate energy consumption prediction. In summary, their proposed hybrid network model, DB-Net, consisting of DCNN and BiLSTM, addresses the challenges of long- and short-term ECP in ILES. The model's preprocessing module ensures optimized data, while the combination of DCNN and BiLSTM enables effective feature encoding and pattern learning. The experimental results demonstrate the superiority of the proposed approach over existing methods, further establishing its effectiveness in energy consumption prediction. [26]

3.1.3 Transformers

Transformer Models: Transformer models, such as the popular architecture known as the "Transformer," have shown promise in energy forecasting tasks. These models excel in capturing dependencies and patterns in sequential data, making them suitable for capturing temporal dynamics in energy time series.

Extending the forecasting time is a crucial requirement in practical applications such as extreme weather early warning and long-term energy consumption planning. This study focuses on addressing the long-term forecasting problem in time series data. Existing

Transformer-based models utilize different self-attention mechanisms to capture long-range dependencies. However, the intricate temporal patterns in long-term forecasts pose challenges in discovering reliable dependencies. Moreover, Transformers often employ sparse versions of point-wise self-attentions to handle long series efficiently, resulting in limited information utilization. To overcome these limitations, we introduce a novel decomposition architecture called Autoformer, which incorporates an Auto-Correlation mechanism. Unlike conventional series decomposition approaches, we integrate it as a fundamental inner block within deep models. This unique design empowers Autoformer with progressive decomposition capabilities, enabling effective handling of complex time series data. Additionally, we leverage insights from stochastic process theory to develop the Auto-Correlation mechanism, which operates at the sub-series level to discover dependencies and aggregate representations based on the series periodicity. Compared to self-attention, Auto-Correlation demonstrates superior efficiency and accuracy. In the context of long-term forecasting, Autoformer achieves state-of-the-art accuracy by outperforming existing models. It exhibits a relative improvement of 38% across six benchmark datasets, covering diverse practical applications such as energy, traffic, economics, weather, and disease forecasting. [27]

Time series forecasting is a crucial task in various domains, including predicting energy output from solar plants, electricity consumption, and traffic congestion. In this research, we address this forecasting challenge using the Transformer model [28]. While we were initially impressed by its performance, we identified two major weaknesses that needed to be addressed: (1) the model's insensitivity to local context, which can lead to anomalies in time series, and (2) the quadratic growth of memory usage in the canonical Transformer architecture, making it impractical for modeling long time series. To overcome these issues, we propose two key enhancements. Firstly, we introduce convolutional self-attention, which incorporates local context by employing causal convolution to generate queries and keys. This modification allows the attention mechanism to better capture local dependencies within the time series. Secondly, we present the LogSparse Transformer, which significantly reduces the memory requirements to $O(L(\log L)^2)$, enabling more accurate forecasting for time series with fine granularity and strong long-term dependencies while operating within constrained memory constraints. We conducted experiments using both synthetic and real-world datasets, and the results demonstrate that our proposed approach outperforms state-of-the-art models in time series forecasting tasks. The LogSparse Transformer exhibits improved performance, offering a more efficient and accurate solution for capturing temporal patterns and predicting future values in time series data. [29]

4 Conclusion

In conclusion, the utilization of data mining, machine learning, and deep learning techniques in energy forecasting offers a wide range of benefits that have the potential to transform the energy industry. These techniques provide a significant improvement in forecasting accuracy, enabling stakeholders to make informed decisions regarding load balancing, resource allocation, and demand response programs. The ability to accurately predict energy consumption patterns empowers energy providers, grid operators, and policymakers to optimize energy generation, storage, and distribution, resulting in cost savings and improved system efficiency.

Furthermore, the real-time adaptability of data mining techniques allows for proactive responses to changing conditions, such as sudden weather changes or electricity price fluctuations. By continuously analyzing and updating forecasts based on evolving factors, grid stability, load management, and energy trading strategies can be optimized, ensuring the reliable and efficient operation of the energy system.

Moreover, the integration of renewable energy sources into the power grid is facilitated by data mining techniques. These methods effectively handle the inherent variability of renewables by analyzing historical renewable generation data, weather patterns, and other relevant factors. Accurate forecasting of renewable energy availability enables grid operators to balance the intermittent nature of renewables, improving grid stability and promoting a cleaner and more sustainable energy mix.

Scalability and efficiency are also significant advantages of data-driven approaches in energy forecasting. With advancements in computing technologies and the availability of large-scale datasets, these techniques can process vast amounts of data efficiently and provide forecasts in near-real time. As energy systems continue to grow in complexity and scale, data-driven approaches offer scalable solutions for forecasting energy demand, optimizing resource allocation, and managing grid operations effectively.

Finally, data mining, machine learning, and deep learning techniques enable continuous improvement over time. These approaches can be trained on new data as it becomes available, allowing for the refinement and enhancement of forecasting accuracy. By incorporating feedback loops and iterative model updates, these techniques can adapt to changing energy consumption patterns, technology advancements, and market dynamics, ensuring the reliability and effectiveness of energy forecasting systems.

References

[1] Wang, Huaizhi, et al. "A review of deep learning for renewable energy forecasting." Energy Conversion and Management 198 (2019): 111799.

[2] Wang, Huaizhi, et al. "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks." IEEE Transactions on Industrial Informatics 14.11 (2018): 4766-4778.

[3] Pillot, Benjamin, et al. "Historical trends in global energy policy and renewable power system issues in Sub-Saharan Africa: The case of solar PV." Energy policy 127 (2019): 113-124.

[4] Zhao, Yongning, et al. "A novel bidirectional mechanism based on time series model for wind power forecasting." Applied energy 177 (2016): 793-803.

[5] Kazemi, Aliyeh, and Mahnaz Hosseinzadeh. "A multi-level fuzzy linear regression model for forecasting industry energy demand of Iran." Procedia-Social and Behavioral Sciences 41 (2012): 342-348.

[6] Ekonomou, Lambros. "Greek long-term energy consumption prediction using artificial neural networks." Energy 35.2 (2010): 512-517.

[7] An, Ning, et al. "Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting." Energy 49 (2013): 279-288.

[8] Kavaklioglu, Kadir, et al. "Modeling and prediction of Turkey's electricity consumption using artificial neural networks." Energy Conversion and Management 50.11 (2009): 2719-2727.

[9] Pappas, S. Sp, et al. "Electricity demand load forecasting of the Hellenic power system using an ARMA model." Electric Power Systems Research 80.3 (2010): 256-264.

[10] Kucukali, Serhat, and Kemal Baris. "Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach." Energy policy 38.5 (2010): 2438-2445.

[11] Vapnik, Vladimir. The nature of statistical learning theory. Springer science & business media, 2013.

[12] Zhong, Hai, et al. "Vector field-based support vector regression for building energy consumption prediction." Applied Energy 242 (2019): 403-414.

[13] Zhang, Zichen, and Wei-Chiang Hong. "Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads." Knowledge-Based Systems 228 (2021): 107297.

[14] Liu, Hui, et al. "Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression." Renewable Energy 143 (2019): 842-854.

[15] Liu, Da, and Kun Sun. "Random forest solar power forecast based on classification optimization." Energy 187 (2019): 115940.

[16] Ahmad, Tanveer, and Huanxin Chen. "Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems." Sustainable Cities and Society 45 (2019): 460-473.

[17] Cai, Jianchao, et al. "Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest." Applied energy 262 (2020): 114566.

[18] Nie, Peng, et al. "Prediction of home energy consumption based on gradient boosting regression tree." Energy Reports 7 (2021): 1246-1255.

[19] Lu, Hongfang, et al. "Short-term prediction of building energy consumption employing an improved extreme gradient boosting model: A case study of an intake tower." Energy 203 (2020): 117756.

[20] Kumari, Pratima, and Durga Toshniwal. "Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance." Journal of Cleaner Production 279 (2021): 123285.

[21] Somu, Nivethitha, Gauthama Raman MR, and Krithi Ramamritham. "A deep learning framework for building energy consumption forecast." Renewable and Sustainable Energy Reviews 137 (2021): 110591.

[22] Jin, Ning, et al. "Highly accurate energy consumption forecasting model based on parallel LSTM neural networks." Advanced Engineering Informatics 51 (2022): 101442.

[23] Wang, Jian Qi, Yu Du, and Jing Wang. "LSTM based long-term energy consumption prediction with periodicity." Energy 197 (2020): 117197.

[24] Sajjad, Muhammad, et al. "A novel CNN-GRU-based hybrid approach for short-term residential load forecasting." Ieee Access 8 (2020): 143759-143768.

[25] Wu, Kuihua, et al. "An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system." International Transactions on Electrical Energy Systems 31.1 (2021): e12637.

[26] Khan, Noman, et al. "DB-Net: A novel dilated CNN based multi-step forecasting model for power consumption in integrated local energy systems." International Journal of Electrical Power & Energy Systems 133 (2021): 107023.

[27] Wu, Haixu, et al. "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting." Advances in Neural Information Processing Systems 34 (2021): 22419-22430.

[28] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).

[29] Li, Shiyang, et al. "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting." Advances in neural information processing systems 32 (2019).