Florida Electricity Prediction Using Machine Learning

Lily Schleider, Qipeng Zheng

EXCEL, DUKE ENERGY, IEMS Labs

UCF The Center for Initiatives in STEM

DUKE ENERGY®

ABSTRACT

Predicting future energy demands will allow for better planning and operation from electricity providers. Suppliers will have an idea of what they need to prepare for, thus preventing over and under estimations. This can save money and make the energy industry more efficient. We apply Convolutional Neural Networks (CNN) in order to predict Florida's future electricity use. The Univariant CNN only accounted for the energy consumption variable. The Multichannel network took into account all the time series variables. The Multihead network created a CNN model for each of the variables that then combined to make a prediction. For all of the models, the dataset was split up into train and test data so the predictions could be compared to the actual values in order to get accuracy. For the CNNs, the Univariant predictions had more diverse and higher standard deviations compared to the Multichannel network and even higher when compared to Multihead network. We can conclude that the Multihead model performed better than the Multichannel model and the Multichannel model performed better than the Univariant model when trying to predict future energy consumption.

DATASET

- The dataset was created from a variety of sources ranging from the United States Energy Information Administration, Florida Climate Center, the United States Census Bureau and the Bureau of Labor Statistics.
- The dataset included 8 years of data (split monthly) ranging from the beginning of 2010 to the end of 2017.
- There are 22 variables as shown in Table 1.

YEAR	Years 2010 to 2018
MONTH	Month 1 to 12 (January to December)
Electricity Demand (MWH)	Electricity Demanded in Florida in MWH
Revenue	Amount of electricity sold in the state in Thousand Dollars
Customers	# of customers buying electricity in Florida
Price	Price of Electricity (Cents/kWh)
Avg Temp	Fahrenheit
Max Temp	Fahrenheit
Min Temp	Fahrenheit
Precipitation	Recorded in Inches
Cooling Degree Days	Fahrenheit degree days
Heating Degree Days	Fahrenheit degree days
GDP Florida	Total Gross Domestic Product for Florida in Millions of Dollars
Population	Population of Florida
Households	# of households
Household Size	Avg persons per household
Labor Force	# of persons in Labor Force
Employment	# of persons Employed
Unemployment	# of persons unemployed
Unemployment Rate	Unemployment rate %
Visitors	# of Visitors traveling to Florida

Table 1: Variables Included in Dataset

METHODOLOGY

- 75% of the dataset (6 years) is used for training and 25% (2 years) is used for testing.
- The model trains with the training dataset and then makes a prediction. The predictions are then tested against the testing dataset in order to see how accurate our model performed.
- Typically, when creating a training dataset, random observations would be chosen from the original dataset. However, in this study, consecutive observations were chosen for the training set since historical data is being utilized and we wanted to account for seasonality in the data.
- The filters in the CNN have random weights. As the filter shifts across the dataset, these weights use matrix multiplication on the data.
- We use the maxpooling layer to keep the dominating features while decreasing the dimensions. Maxpooling keeps the highest value and gets rid of everything else.
- In order to measure accuracy, the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2) between the actual and predicted values were calculated.

Univariant CNN

- The Univariant CNN only uses the electricity demand data in order to train the model.
- There are two 1D convolutional layers that each had 64 filters and a kernel size of 2x1. The ReLU activation function is used.
- The 1D maxpooling layer had a size 2x1.
- The model is flattened and then goes through a Long-Short-Term-Memory (LSTM) network.
- Lastly, it goes through time distributed dense layers.
- The CNN had 1000 epochs and a batch size of 24.

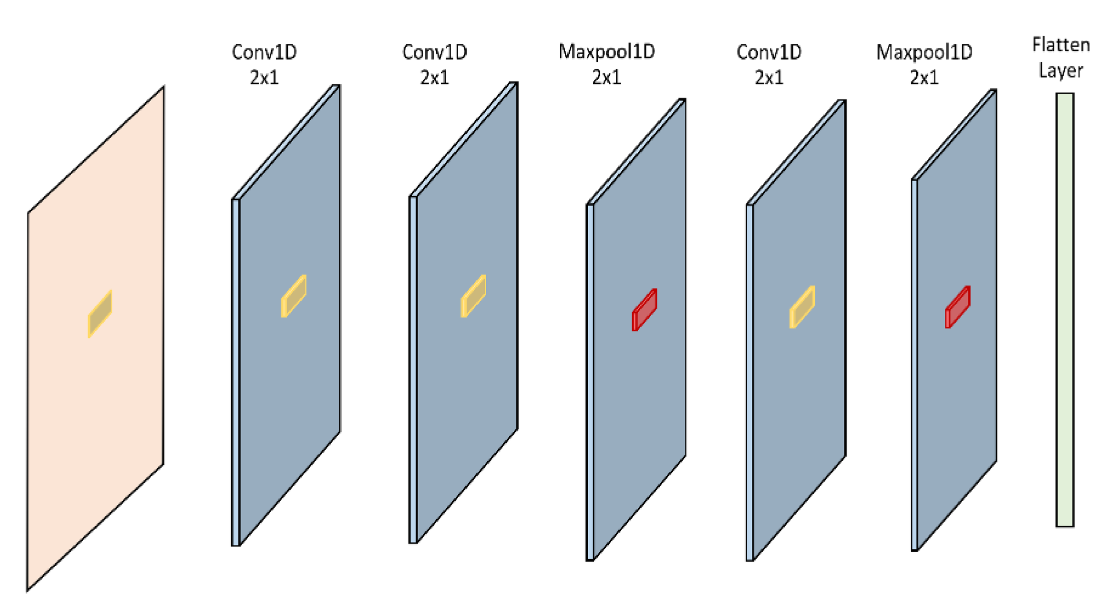

Figure 1: Convolution and Maxpooling Layers for Univariant CNN

Multichannel

- This CNN differs in the additional use of all the time series variables.
- We set up each time series variable as its own channel of input. This gives our model more information to work with and works particularly well when the output is a function of the inputs. (Ahangar)
- The CNN has three 1D convolutional layers that have 64 filters, 64 filters, and 16 filters respectively. They each have a size of 2x1. The ReLU activation function is used.
- There are two 1D maxpooling layers with a size 2x1.
- There is a flattening layer followed by fully connected dense layers.

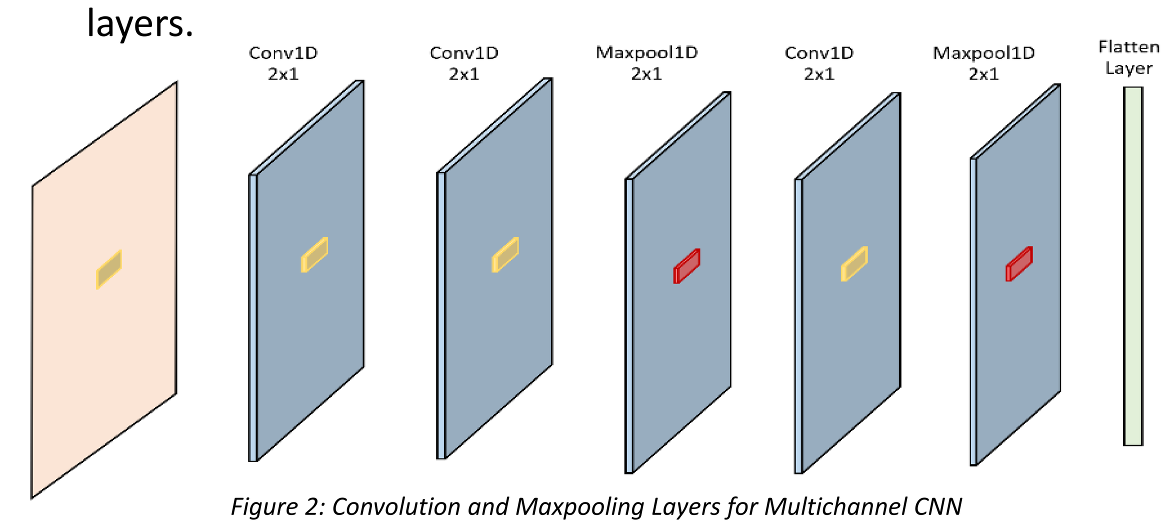

Figure 2: Convolution and Maxpooling Layers for Multichannel CNN

Multihead

- The Multihead CNN uses a sub-CNN model for each input variable. For each time series variable, we take the 1D input that has n inputs and put it through a model that outputs a flat vector. This flat vector summarizes the features of the sequence. We can combine all the outputs for each variable through concatenation.
- Each model has three 1D convolutional layers that each has 64 filters, 64 filters, and 16 filters respectively. They each have a size of 2x1 and ReLU activation.
- There are two 1D maxpooling layers that had a size of 2x1.
- There is a flattening layer. All the flattened layers are concatenated together and go through fully connected dense layers.

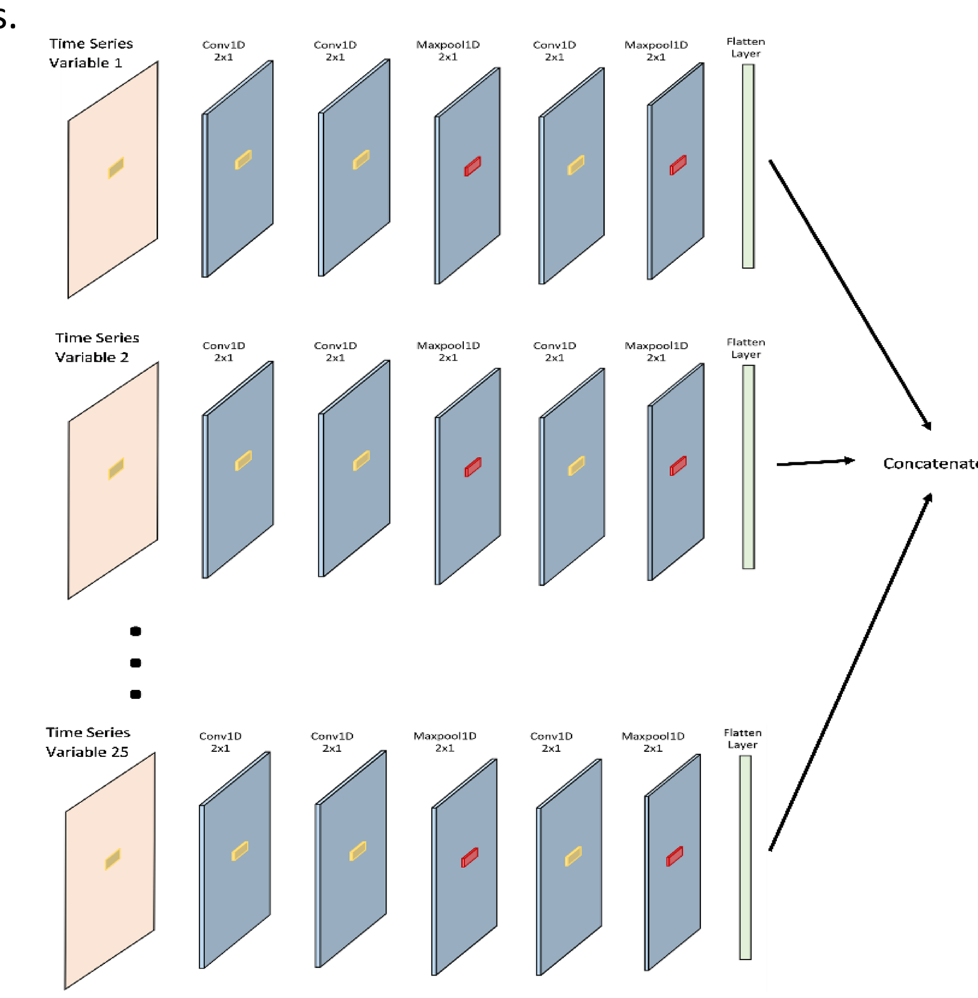

Figure 3: Convolution and Maxpooling Layers for Multihead CNN

CONCLUSION AND FUTURE WORK

- The Mulitchannel CNN did a much better job at predicting the next 24 months compared to the Multihead and Univariant CNN.
- The Multichannel was also much better at predicting abnormal raises in the dataset.
- In the summer months of the first year predicted, all the models underpredicted. The years before had lower energy consumption and the years we predicted had an unusual demand spike, which the models didn't expect.
- Being able to predict future electricity demand will improve planning from electricity providers and suppliers. They will be able to plan ahead to create a more efficient strategy which will save money and prevent over and under preparation.
- For future work, we would like to train the models on a larger dataset in case lack of data was contributing to the results. We would also like to test these models on other non-linear datasets.
- Figure 4 shows the last two years of electricity demand along with the predictions made by the different methods.

	Univariant	Multichannel	Multihead
MSE	8302957215341.45	342181554643.63	450701238904.58
RMSE	2881485.24	584962.87	671342.86
MAE	2308164.96	474780.29	466385.17
MAPE	0.1148767	0.0250089	0.0231797
R^2	0.0705	0.963100	0.9498

Table 2: Model Performance Comparison

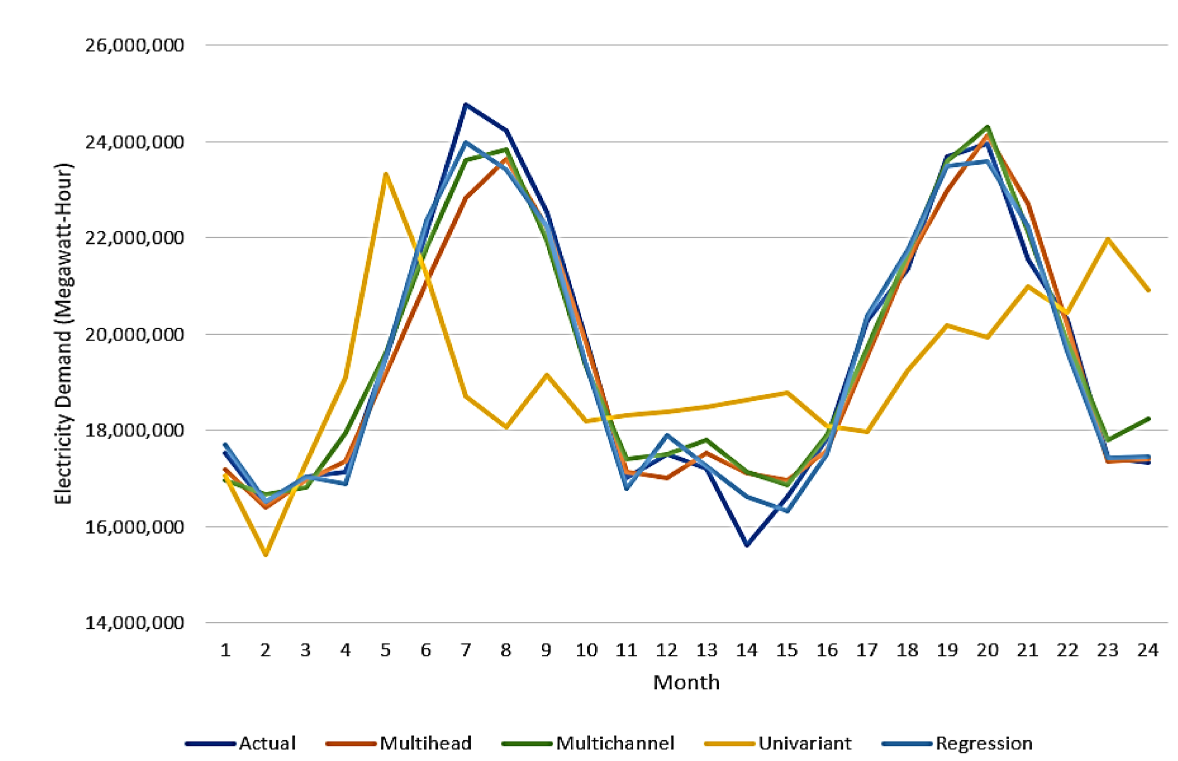

Figure 4: Actual vs Predicted Electricity Demand for All Models

REFERENCES

Ahangar, R.G., Yahyazadehfar, M., & Pournaghshband, H. (2010). The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for Prediction Stock Price in Tehran Stock Exchange. *ArXiv, abs/1003.1457*.

Hamedmoghadam, H., Joorabloo, N., & Jalili, M. (2018). Australia's long-term electricity demand forecasting using deep neural networks. *ArXiv, abs/1801.02148*.

Lee, K. Y., Kim, K. H., Kang, J. J., Choi, S. J., Im, Y. S., Lee, Y. D., & Lim, Y. S. (2017). Comparison and analysis of linear regression and artificial neural network. International Journal of Applied Engineering Research, 12(20), 9820-9825.

Tsakiri, Katerina & Marsellos, Antonios & Kapetanakis, Stelios. (2018). Artificial Neural Network and Multiple Linear Regression for Flood Prediction in Mohawk River, New York. Water. 10. 1158. 10.3390/w10091158.