**Introduction**

In my Data Science project I want to forecast the number of unit sales per month by product name and by store number. In order to explore the ultimate model of prediction I used a dataset of Favorita stores located in Ecuador. There is a big motivation from groceries to have a model which forecasts the sales accurately. It comes from the big need of not being stuck with overstocked, perishable goods and products which finished their self-life and need to be destroyed. More forecasting systems are based on a statistical learning method like moving average, and regression which are reasonably effective in most cases but require heavy and delicate feature engineering. In the past, in order to predict unit sales Alexia Wenxin Xu, a researcher from Stanford proposed to use a residual neural network to output a predicted unit sales. She recognized that both neural and residual networks improved the model by decrease the dev loss but there was a big noise in the dataset (1). Another group tried to solve the challenge of forecasting the product sales by applying the convolutional neural network method in the context of time series data. (2) They also used the Favorita stores Dataset and found out that convolutional neural networks worked good at handling historical data and catching seasonality, trends, cycles, and irregular components. Using a CNN WaveNet, a sequence to sequence architecture in order to do sales forecasting was a very effective method of solving time series predicting problems. (3) Another group tried to combine between the deep learning method and the machine learning method by using LSTM and combine it with lightGBM, the first one used in order to dig in the data and get information from while the second used for being with strong interpretability. The combination between these two models used to forecast the supply chain sales by comparing the implementation of other different models and discuss the advantages of the combined model as well. Similar to first two projects which are discussed above they also took the Favorita Grocery’s supply chain sales data set for running the combined model on. Their insight was that the combined model predicts the supply chain sales accurately and the interpretability of the predicted results was significant to improve many aspects such as production mode, price management and precise marketing of supply chain. In my project the outcome feature is defined by the total item number sales per month, by store number.

1. <https://pdfs.semanticscholar.org/1b5f/d7be136793a213b91769344bdb4c72b99fe9.pdf>
2. <https://arxiv.org/pdf/1803.04037.pdf>
3. <https://www.emerald.com/insight/content/doi/10.1108/IMDS-03-2019-0170/full/html?casa_token=Qs4fCrj6seEAAAAA:ZbkIoCtbQiU727yeZv2ZQ00U87aRDNPNXeQuNTOCtaDVWAF69k-OWVSsB3IOkaBP4c3PsmH7K4Ul_1ag9zeQ3tOFnRtoNlSuTj21RcloBMUDGWa4ERJH>