MIDDLE TENNESSEE STATE UNIVERSITY

Demand Forecasting by use of Long-Short Term Memory (LSTM) Architecture

Term Project – Deep Learning 7850

By Richard Hoehn

MTSU – November 30th, 2023

Welcome

Richard Hoehn

•

PhD Student – Computational & Data Sciences

•

Living Franklin

•

MTSU Grad 2005, Vanderbilt Grad 2009

•

Last 15 years in SW development for Retail, Mfg. & Logistics Applications

•

Primarily in Databases (PostgreSQL), API / Pub-Sub (Queues), and Business Reporting Apps

Introduction & Agenda

Introduction & I will Speak About

- Significance of Demand Forecasting and Prediction Results
- Motivation for using LSTM Models on Historical Sales Dataset
- Key Goals & Aims for Term Project -> The Methodology, Process, and Results
- *and finally,* Presentation of Visual Results for Review

Agenda

- Significance of Demand Forecasting & Background, Literature Review
- Motivation and Scope of this project including Key Goals & Aims
- Methodology – Details on process flow and Completed Work
- *and finally,* Analysis of Results, Conclusion, and Future Work

Significance of Demand Forecasting

Types

Qualitative Forecasts

- **Expert Opinions:** This involves gathering insights from individuals within or outside the organization
- **Market Research:** Involves collecting data through surveys, focus groups, or customer interviews to predict future demand.

Quantitative Forecasts

- **Time Series Analysis:** Utilizes historical sales data to identify trends, patterns, and seasonal variations.
- **Causal Models:** These models believe that demand is influenced by certain factors or events.

Market

By accurately identifying and understanding Sales Demand companies can:

- Optimize Inventory & Production
- Manage Supply Chains (Re-Supply)
- and help with Financial Planning.

The market for Demand Forecast software and services is estimated to be **$3.62billion**[1] in **2022.**

LSTM Architecture

Motivation for Demand Forecast Project

- Time-series problems are interesting!
- I often question how one can predict the future based on past data.
- Additionally, how the future might change if the input of data had changed?
- A better understanding of RNNs in specific the LSTM architecture will better my understanding of this deep learning field.
- This will prepare me for my future research in Computation & Data Sciences for forecasting model design.

Objectives & Key Aims

1. Deploy a dataset in the cloud (AWS) for remote extraction, process, and normalization that can then be used for training and validation.
2. Build, Train, and Validate the LSTM model with the pre-processed dataset from cloud.
3. Using plotting techniques provide a visual approach to reviewing performance of the LSTM model.

In summary, the **student learning** was how to **P**rocess, **B**uild, and **T**est the **LSTM** model for sales demand prediction using a publicly available dataset.

•

and... **V**isualize the validity of using the LSTM architecture for sales demand forecasting.

Literature Review

- Unlike traditional RNN architectures the LSTM design can regulate the forward and backwards flow of information and more importantly be able to forget non-valuable data-traits[4].

***”LSTM are adept at reading Seasonal Traits in Demand Forecasting.”* [3]**

- LSTM models have emerged as a ”go-to” tool[4, 5] in the field of demand forecasting.

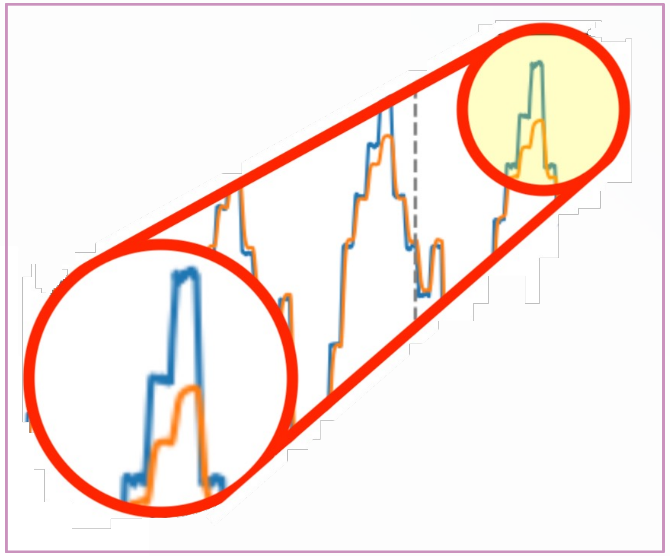
Methodology

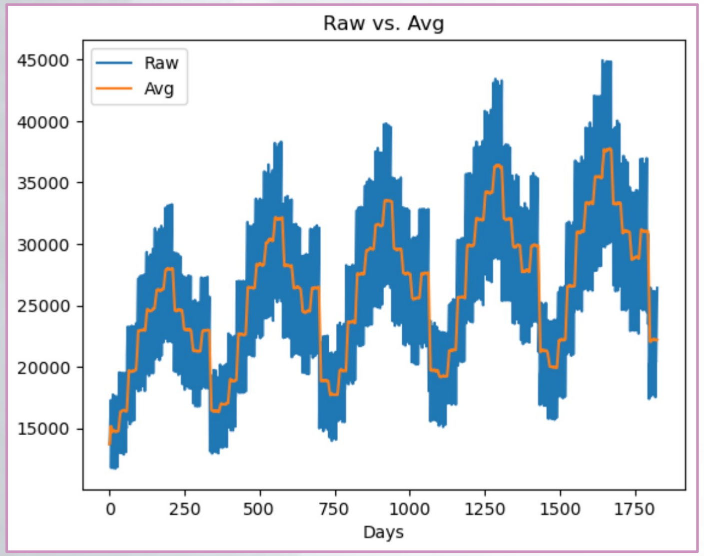
Data Procurement

Parsing & Cleanup, Normalization

LSTM Training & Validation

Visual Review





Methodology

Data Procurement & Review

- The dataset is from Kaggle[2] demand forecast challenge of about five (5) years ago.
- It is comprised of five (5) years of store sales data.
- Based on 50 different items at 10 different stores.
- Each data point is a single day's data-point (rows) with this being said the dataset is comprised of 912,500.
- In order to make this a "Uni-Variante" project I grouped that all the sales by "date".
- Equation 1 – Depicts the Group-Summation by <u>Date (Day)</u>

Dataset Details

Years	5	Days	1,826
Stores	10	Items per Store	50
Datapoints	912,500	File Type	csv
Min ($USD)	11,709	Max ($USD)	44,936
Mean	912	Std.	527

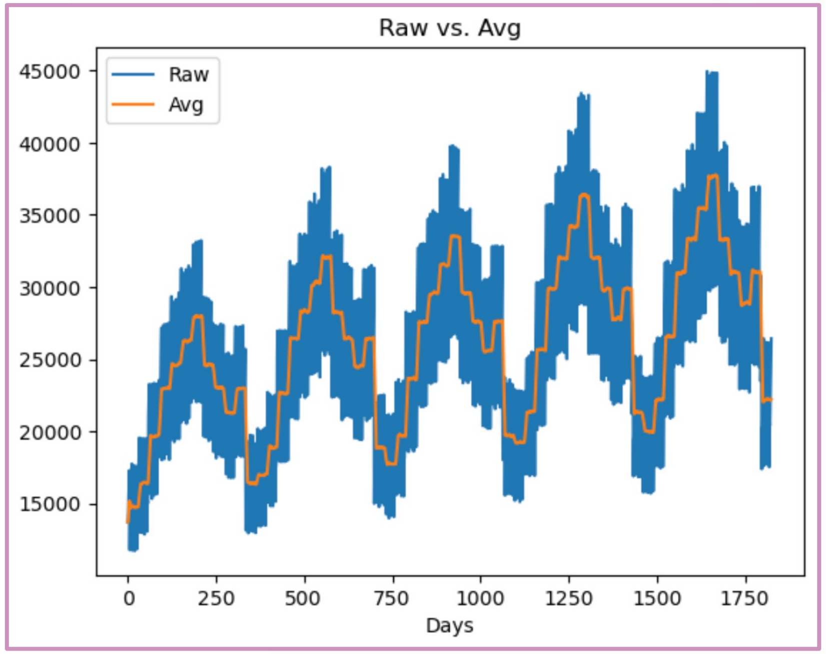
$$s_i = \sum_{j=1}^{n} a_{ij} \tag{1}$$

In this Equation 1 the following is given:

- i index for days
- j index for stores
- aij total sales of store j on day i
- si total sales across all stores on day i

Methodology
Data Smoothing & Normalization

- The dataset is derived from real-world sources, suggesting it's inherent noise.
- I used a smoothing algorithm to help the model learn more easily without the daily noise. A good methodical approach was to use a **7day** moving average approach[Eq:2].
- I chose for this project was using the Min-Max Scaler[Eq:3], where the minimum value of the data becomes **-1** and the maximum value becomes **1**.



$$\mathrm{MA}_i = \frac{1}{7}\sum_{k=i-6}^{i} s_k \qquad (2)$$

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (3)$$

Methodology

Data Sequence – 90days & Setup

- After smoothing and normalization, I divided the data into overlapping 90-day.
- Example, the first sequence consists of sales data from January 1 to March 31, and then the next from January 2 to April 1, and so on. [Eq:4].
- Once X & Y datasets where created I split them into the **80 / 20** for training and validation.
- I choose 90days to capture the seasons (4x) per year.

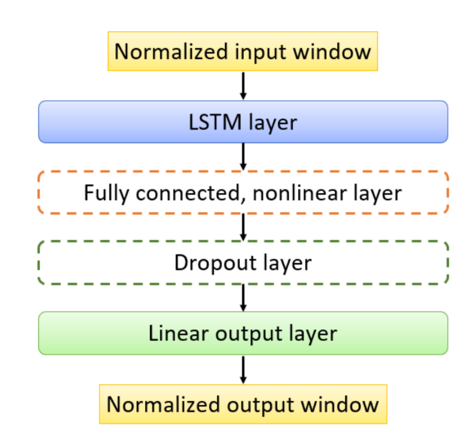
$$\mathbf{X}_i = \begin{bmatrix} S_{t-89} \\ S_{t-88} \\ \vdots \\ S_t \end{bmatrix} \quad (4)$$

```
# Function to create sequences
def create_sequence(data, seq_length):
    xs = []
    ys = []

    for i in range(len(data) - seq_length):
        x = data[i:(i + seq_length)]
        y = data[i + seq_length]
        xs.append(x)
        ys.append(y)

    return np.array(xs), np.array(ys)

# Create sequences
x, y = create_sequence(df.values, 90)
```

Methodology
LSTM Model Build & Training

The training time for this compact model was brief, **~6** minutes on the HPC GPUs performing 150 epochs.

Mean Squared Error (MSE) loss is commonly[3] used as a learning metric, quantifying the difference between predicted and actual values.



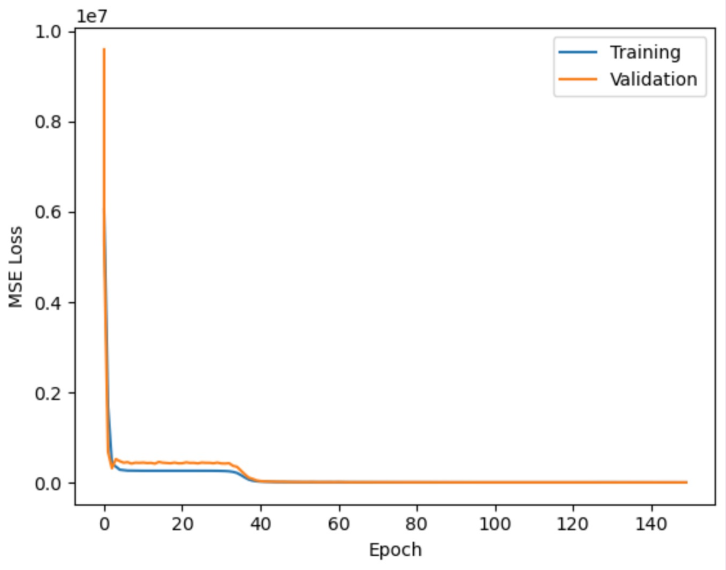
$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

LSTM Parameters

Learning Rate	0.001	Dropout	0.1
Hidden Layers	50	LSTM Layers	4
Number of Epochs	150	Sequence Size	90d

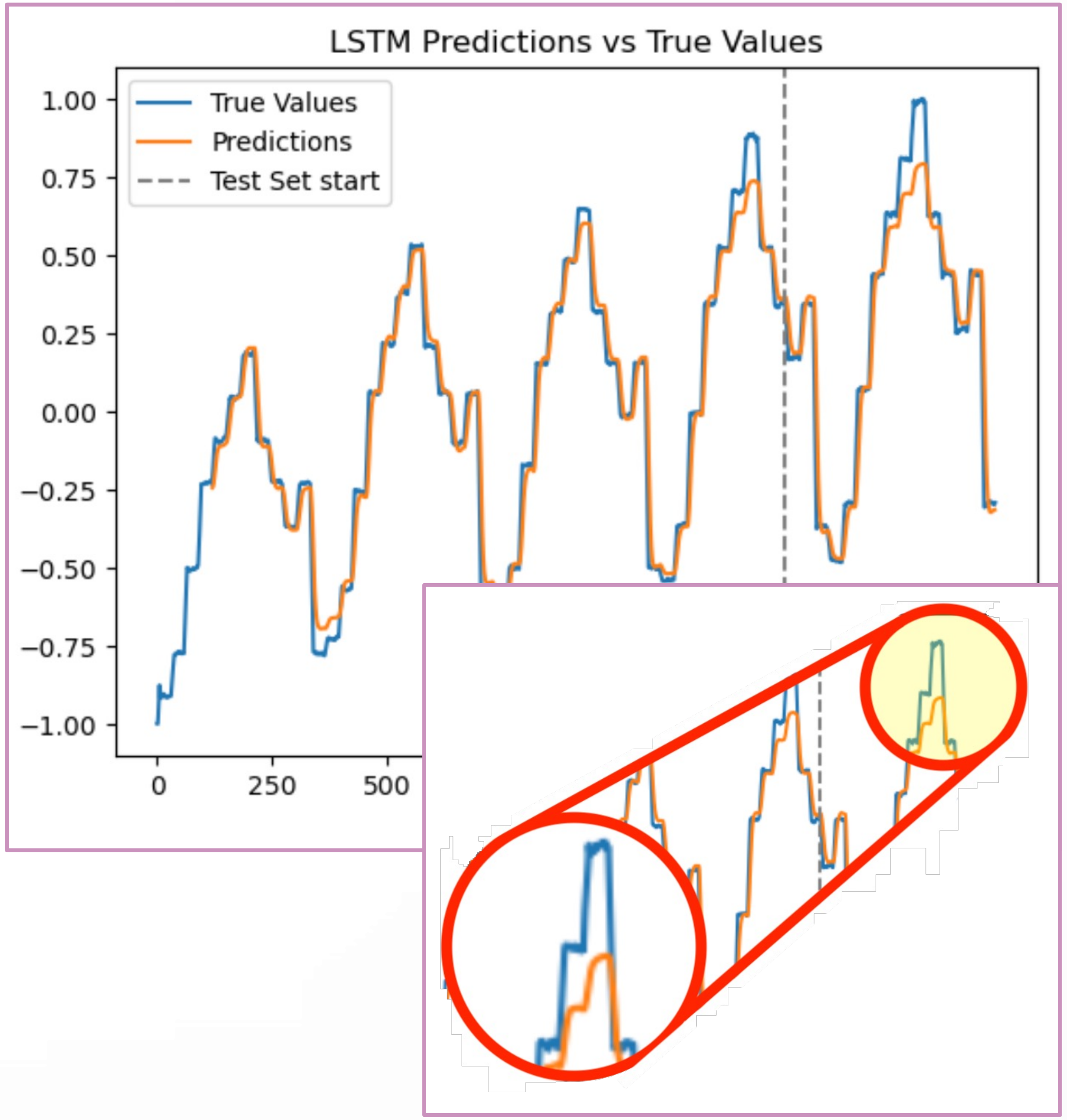
PyTorch

```python
class LSTM(nn.Module):
    def __init__(self, input_size, hidden_size, num_layers, output_size):
        super(LSTM, self).__init__()
        self.hidden_size = hidden_size
        self.num_layers = num_layers
        self.lstm = nn.LSTM(input_size,
                            hidden_size,
                            num_layers,
                            dropout=0.1
                            batch_first=True)
        self.lin = nn.Linear(hidden_size, output_size)

    def forward(self, x):
        y = x
        y, _ = self.lstm(y)
        y = self.lin(y[:, -1, :])
        return y
```



Results & Analysis

- The training validation of the model are quite promising. The visual to the right illustrate that it overall follows the true path during the training and test phases.
- In the main visual it is apparent that the LSTM model as able to learn the underlying seasonality of the sales and would therefore be able to predict future demands.
- As one can see in the detail visual it seems as time moves forward that the difference between true vs. predicted values skews higher over time.



Conclusions & Future Work

Concluding

The LSTM model worked well for learning demand over time and recognized seasonality shifts in the sales demand.

As time passed the model did start failing to predict the “highes” of the season. The reason for this is not fully apparent and will require further analysis.

Areas of Future Work

- **First,** I intend to explore the extraction of additional contextual information related to the timestamps, such as seasons, weekdays, holidays, or any other relevant time-based factors.
- **Second,** with additional information the problem becomes a “multi-variante”. Comparing to the non-expanded dataset presented previously analysis might shed light on the extent to which the inclusion of more temporal information influences forecasting accuracy.
- **Third,** find a way to show the accuracy of the model in dollars so a business could use ti to better predict needs in the future.

Thank you

Citations

1. Zion Market Research – Demand Planning Solutions Market. URL: https://www.zionmarketresearch.com/report/demand-planning-solutions-market
2. Kaggle. Store Item Demand Forecasting Challenge. 2018. URL: https://kaggle.com/competitions/demand-forecasting-kernels-only
3. Tong Zhou. “Improved Sales Forecasting using Trend and Seasonality Decomposition with Light-GBM”. In: 2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD). IEEE, May 2023. DOI: 10 . 1109 / icaibd57115 . 2023 . 10206380. URL: http://dx.doi.org/10.1109/ ICAIBD57115.2023.10206380.
4. Barıs¸Karaman. Predicting Sales, Forecasting the monthly sales with LSTM. URL: https://towardsdatascience.com/predicting-sales-611cb5a252de
5. Abeselomgebrekidan. Pharmaceutical Sales predic- tion Using LSTM Recurrent Neural Network. URL: https://medium.com/@abeselomgebrekidan12/pharmaceutical-sales-prediction-using-lstm-recurrent-neural-network-db0f980572cc