AISHWARYA SWETHA JONNALAGADDA

857-544-8161| jonnalagadda.ai@northeasern.edu | https://www.linkedin.com/in/aishwaryaswetha/|
Boston, MA| https://github.com/AishwaryaSwetha

EDUCATION

Northeastern University, Boston, US — Sep 2019-May 2021
Candidate for Master of Science (MS) in ***Data Analytics and Data Science*** — **GPA**: 4

Mahatma Gandhi Institute of Technology, Hyderabad, India — Aug 2014 -Mar 2018
Bachelor of Technology in ***Computer Science and Engineering***
Received outstanding ***Academic Excellence*** *award from JNTU-H*

TECHNICAL SKILLS

- **Programming languages**: C, C++, Java, Python, R, MATLAB
- **RDBMS/NOSQL**: MySQL, PostgreSQL, Oracle dB, Neo4j, MongoDB, Bigtable, HBase, Dynamo DB
- **Packages/Framework**: Keras, TensorFlow, PyTorch, Scikit-learn, Spark, Django, Kubernetes, Docker, NumPy, Pandas
- **Tools**: Jupyter Notebook, R studio, Microsoft Office, GitHub, Google Colab, Tableau, Amazon Web Services, Google Cloud Platform, Amazon EC2, Amazon Sage Maker, Flask

PROJECTS

Deep Learning using PyTorch — **Summer 2021**

- Participated in a 6-Week course of Deep Learning with PyTorch, which includes coding, hackathon, Assignment each week.
- Implemented PyTorch Torch components, Melanoma Classification project using Resnet-50 with 54% accuracy.
- Protein Classification problem from Kaggle competition stood among top 15% of competitors.

Stock Price Prediction using Artificial Recurrent Neural Network -LSTM of Corporation Summer 2021

- Performed Time Series Analysis on Stock price attribute, namely 'Close' and Analyzed the trendline.
- Built LSTM Neural Network using Keras package and predicted stock price the following week or month accordingly.
- Model on the test data/new data has lower error rate of 5.31 (RMSE)Root Mean Square Error and hyperparameters like 80:20 split of dataset,10-fold cross validations on the splits are used to evaluate and choose the best model performance on training.

Image Classification on CIFAR-10 using Tensor Flow — Summer 2021

- Used CS-231n Stanford course instructions, built KNN, SoftMax classifier and SVM on CIFAR-10 dataset.
- Performed Logistic Regression, Random Forest, Convolutional Neural Network (CNN), with 4-layers CNN and 6-layer CNN.
- Accuracy of models have been increased from 38% (logistic regression) to 90% (CNN with n layer), CNN gives the best model performance.

Big Mart Sales Prediction using Regression Models-Kaggle — Spring 2021

- Performed Exploratory Data analysis, Data Preprocessing, Feature Engineering, Feature Transformation on Big mart Sales data product wise and sales wise parameters reducing and manipulating 60% of overall inconsistent data.
- Applied Principal Component Analysis to find most contributors on Item_Outlet_Sales which was our Target Variable.
- Evaluated Accuracy using Lift Charts, Evaluation Metrics such as RMSE and compared each model fit, we improved our model performance by 30% by using XGBoost Model which was suitable for our dataset.

Fake News Detection Using NLP — Spring 2021

- Performed Text pre-processing techniques and NLP Features such as (TF-iDF, Bag of Words) is applied on data.
- Applied Logistic Regression on the data and required accuracy measure has been applied.
- Model is deployed using Flask components.

EXPERIENCE

Predictive Analyst -ANZ Virtual Internship — May 2020-Jun 2020

- Performed Exploratory Data Analysis using Python on synthesized transaction dataset of three months transactions for 100 customers, Data Cleaning and Feature Engineering is performed with about 76% of improvement in data quality for model fitting.
- Insights have been made such as average transaction amount each month, spending's of customers per week, Outlier Analysis which distorts the analysis, visualizations have been made and documented.
- Implemented Predictive model suitable for the dataset, fit linear regression which gave poor results which 0.37% R-square hyperparameters such as learning rate, number of features have been tuned with 0.45% R-square.
- Executed Decision Tree on this model for salary prediction which gave .78% R-square and correlation between variables have been plotted.

Software Engineer -Automatic Data Processing (ADP), IND — Mar 2018-Aug 2019

- Created self-service Analytical Dashboards which could help end user to investigate their payroll distributions
- Developed Python Scripts which used beautiful soup frameworks for evaluating web-scraping models
- Designed ad-hoc SQL queries for accessing and manipulating data on MySQL
- Performed Predictive Analytics on ADP workforce data predicting number of clients per business unit using Random Forest
- Analyzed Trends such as Overtime, Turnover, etc. on closely working with ADP Data Cloud Team
- Rewarded and Appreciated for being *Best Innovative Employee* among 20 chosen people under SBS-Business Unit