ASHIS PANDA

DATA SCIENTIST

ashisk.akp@gmail.com
medium.com/@GeneAshis
+91-7358270790
ashiskumarpanda
ashis-panda-b211b592/
CaptainAshis

Summary

Data Scientist with 5+ years of Experience executing data-driven solution to increase efficiency and accuracy. Experienced in creating machine learning models using predictive data modelling techniques and analyzing the output of the algorithm to deliver insights and implement action oriented solutions to complex business problems.

Skills

- PYTHON
- SQL
- R
- DATA SCIENCE
- MACHINE LEARNING ALGORITHMS
- TABLEAU
- PYTORCH
- KERAS
- MATPLOTLIB
- NUMPY
- PANDAS
- DEEP LEARNING
- SCIKIT LEARN
- AWS SAGEMAKER
- NATURAL LANGUAGE PROCESSING
- TRANSFORMERS
- LSTM
- PYSPARK
- MULTIPROCESSING

Education

Bachelor Of Technology in Electrical and Electronics Engineering at Vellore Institute of Technology , India — 2011 to 2015

Employment

Abzooba
Data Scientist — July 2020 to Current

G2 — Bangalore
Data Science Analyst — Nov. 2019 to July 2020
This position was a part of Covid19 layoff at G2.

Infosys
Senior Analyst — Dec. 2015 to Oct. 2019

Projects

1. Credit Risk Model Probability of Default (NPA Prediction) -

Probability of the default Model is about describing the likelihood of a default over a particular time horizon. It provides a probability estimate that a borrower will be unable to meet its debt obligations.
Worked on business requirement/model document for development of propensity model for Probability of Default.
i). This will be responsible for providing analytics support across credit risk lifestyle initiatives: origination, portfolio/capital & customer risk management with collections & recovery.
ii). Worked on Statistical models Logistic Regression /Decision Trees in JupyterLab platform for scoring and batch processing using Python Scikit learn, along with the performance of models.
iii). Compared & Selected best model based on ROC Curve & various other parameters like K-S statistics, Lift and Gain Charts.

2. Predictive maintenance using machine learning models

i). Predict possible/upcoming failure in manufacturing equipment. Detect anomalous behavior through sensor readings.
ii). Pre-emptive action based on anomalous behavior to avoid breakdowns.
iii). Saving in terms of both cost and effort.
iv). Reduced downtime and increased efficiency.

3. Moderation Automation algorithm

Used Natural Language Processing on reviews to build G2's very first Moderation Automation algorithm to increase quality and efficiency of review QA . It helped in reducing Manpower hour involved in the process . Processes includes:-
i). Used text analytics to detect Plagiarised reviews.
ii). Implemented Multi-lingual Transformers(B.E.R.T) to detect cross-Language plagiarised reviews.
iii). Conceptualized an algorithm to detect Gibberish reviews present in the database.
iv). Automated the whole of Moderation Automation process .

4. Developed an efficient Dynamic pricing model for one of the largest Logistics Company in USA

This project involved one of the Largest Freight company in USA who acts as a broker between their Clients who wants their loads to be transported and the Carriers.
The Freight Company would get a quotation price by their Clients for the load to be transported between a Origin to Destination pair and would negotiate that price (Carrier Paid Transportation Expense) with the Carriers. This was basically done by the Carrier Representatives in the Freight company by using their existing domain knowledge and their years of experience in negotiation.
Our main aim was to develop an efficient dynamic pricing model which understands that given the factors involved, it can exactly come up with a price that has to be paid to the Carriers. Responsibilities undertaken as a part of this project :-
i) Developed the conceptual and analytical framework of the proposal, as well as the codes for the algorithm used for analysis.
ii) Implemented various ML model for all the lanes and the data was sequenced temporally.
iii) Used Random Forest , XgBoost , Ridge and Lasso models and took an ensemble approach on top of these models, to come up with the best and efficient Carrier Paid Transportation Expense model.
iv) Checked the performance separately on Smooth and Non-Smooth Lanes.

5. Recommendation System

Developed a Recommendation system using Neural Network.
i) A recommendation system, in general consists of a User and an Item.
ii) Each user and each Item were randomly initialized with a set of random numbers also known as Entity Embedding. These embeddings were passed as an input to a Neural Network and it gets updated as we train our Neural Network.
iii) At the end of our training, it's found that the User and Product embedding which were random values initially, has been updated during the training phase and these values starts making sense. These updated set of numbers represents the characteristics of Product and the Preference of Users.
iv) Using these updated Entity Embedding a Recommendation system was developed.

Activities

Towards Data Science and Hackernoon publication. · Writer — 2018 to Current

- I work with Towards Data Science and Hackernoon publication as a Technical Blogger. These are the biggest Machine Learning publication with an outreach of more than 100k subscribers.
- A couple of my blogs have received good appreciation from eminent deep learning personalities like Jeremy Howard (Deep Learning Researcher , President of Kaggle) and David Robinson (Chief Data Scientist at Datacamp) .
- Achieved Rank 296/3714 on the Private Leaderboard in the Analytics vidhya hackathon.

Yes Bank · Guest Blogger
YesBank selected my blog to cover about their mega Datathon event held at Bangalore.

Analytics Society Meetup
I am part of Analytics Society Meetup at my workplace. My responsibilities involve brainstorming concepts of Deep Learning and making the participants understand my experience of solving a project using Deep Learning approach.