Punya Swaroop Sirigiri

959-200-1546 • punyaswaroop12@gmail.com • LinkedIn • GitHub

AWS Certified Cloud Practitioner

Summary: Machine Learning engineer with strong math background and 5+ years of experience using Machine Learning and Data Engineering to solve challenging data science problems in various industries

Achievements: Insurance Excellence Synergy Team Award (John Hancock) from **CEO**, Academic Excellence Scholarship (UCONN), Walmart Labs Hackathon Finalist (Sponsored Event), **Kaggle - (Top 9%, 14%),** Bloomberg Market Concepts

Professional Experience:

Progressive Insurance (Cleveland, OH): **Feb 2020 - Present**

Senior Data Scientist/ Machine Learning Engineer (Python, R, SQL, AWS, Bash, Jenkins, Docker, GIT, Metar API)

- Built, refined and enhanced **Time series forecasting models** (SARIMAX, Prophet etc.,) to predict prospective frequency of automotive accidents and average cost to repair across 50 states, 2 distribution channels and 10+ policy coverage segments (2000+ models) to inform rate-making decisions
- Defined the rate-making response to changes in driving behavior and economic conditions due to COVID19 (lockdowns, WFT, Supply chain disruptions etc.,) resulting in model enhancements yielding sustained growth, profitability and an aligned response across business units (pricing, claims, control, product, R&D)
- Collaborated on the development of an internally facing **Web application** providing interactive tools for 100+ users to meaningfully build intuition about our time series models and consume output to drive business decisions
- Led efforts pertaining to development of **Combined Ratio Probabilistic Model**, which empowered business partners to understand their probability of attaining yearly profitability targets to help with strategic planning and marketing

John Hancock Financial Services (Boston, MA): **Feb 2018 – Feb 2020**

Data Scientist – Insurance and Global Wealth & Asset Management **(Python, SQL, Hive, D3.js)**

- Implemented Random Forest & Logistic Regression to predict whether Financial Advisor would bring in new business in next 12 months using historical retirement plan data (Proposals, Contracts, AUM) for past 10 years, investments data, industry data(external) with best model of lift 5
- Developed Logistic Regression to rank the potential customers based on likelihood to qualify for insurance, click on ads etc., for direct-to-consumer marketing campaigns using customer data, health related information and web analytics data, evaluated performance of models using data of existing customers' data.

Machine Learning Projects:

Conversational Chatbot (NLP)

- Trained a sequence-to-sequence deep learning model on my social media conversation logs to create a chatbot that talks like me. Deployed the trained model to a server using the Flask framework and hosted using Heroku

Zillow's Home Value Prediction (Kaggle - Evaluation Metric)

- Performed explanatory analysis, feature engineering and built simple stacking of LightGBM, CatBoost models to predict the log error between Zestimate and the actual sale price of houses with MAE of 0.07517(Top 10%)

Image Captioning (Computer Vision, CNN)

- Built encoder-decoder model with attention mechanism to generate a descriptive caption for an image we provide it with BLEU-4 score of 33.29

Education:

The University of Connecticut, Hartford, CT Dec 2017
Master of Science in Data Science (Business Analytics) **GPA: 3.80/4.00**
National Institute of Technology, Calicut May 2016
Bachelor of Technology in Engineering **GPA: 8.12/10.00**

Skills:

Machine Learning and Statistics: ARIMA, Linear, Ridge, Lasso, Logistic regression, Decision Trees, Random forest, SVM, Clustering, PCA, KNN, XgBoost, Hypothesis Testing, Probability Distributions, Bayesian Statistics

Programming Languages and Software: Python (Flask, scikit-learn, numpy, pandas, matplotlib), R Programming, Bash, Jenkins, Docker, GIT, SQL, Hive, JavaScript (D3.js), HTML, CSS, Tableau, Excel

Deep Learning and NLP: TensorFlow, Keras, PyTorch, Spacy, OpenCV, Convolutional Neural Networks (CNN), Sentiment Analysis, Sequence Models (RNN, LSTM), Word2Vec, N-grams, TF-IDF, SVD, NLTK

System design and Architecture: AWS S3, AWS EC2, Load Balancing, Networks and Protocols, Terraform, Databases, REST APIs, Consistent Hashing, Caching, Unit testing and