AKSHAY TIWARI

(417)-450-2517 | akkiit.tiwari@gmail.com | San Francisco | LinkedIn /akshaytiwari19 | Github /akkiittiwari | akshay.theia-ai.com

EDUCATION

Master's in Data Science, University of San Francisco, San Francisco   *July '17 – June '18*
- Coursework in machine learning, deep learning, applied statistics, A/B testing, SQL, distributed computing

Bachelor's in Computer Science, University of Delhi, Delhi   *July '12 – May '15*
- Coursework in C++, data structures, database management, algorithms, statistics and probability

EXPERIENCE

Flyr, San Francisco, CA   *October '17 – June '18*
Data Scientist Intern *[Python(pandas, numpy, scikit-learn), SQL(Google BigQuery), AWS(EC2, S3)]*

- Built Machine Learning models to predict the probability of ticketing and cancellation using real-time airline ticketing datasets from industry leaders and achieved AUC scores of 0.77 and 0.71
- Extracted and transformed data stored in BigQuery to use for model training and processed the output of these models as an input to pricing and synthetic inventory engine in Revenue Management architecture
- Delivered a proof of concept based on Random Forest and Linear Regression for predicting the customer travel budget to deliver targeted and optimally priced travel packages
- Analysed past customer data to estimate net wealth and segment customers based on price sensitivity using clustering techniques to personalize and improve the predictions from pricing engine

Inbredia Technologies, Delhi-NCR   *July '15 – Nov '16*
Data Analyst *[Node.js, SQL, MongoDB]*

- Analysed the personnel data for a physical security company to predict and reduce the attrition, delivering a 3% (~1,500K INR) increase in the profit margin
- Developed backend APIs using Node.js and MongoDB for an e-commerce store analytics dashboard projecting business specific metrics to the admin
- Analysed past daily ridership data to come up with an on-ground marketing campaign structure for newly launched routes & implemented it through leading a team of marketing executives for a last-mile transport company

PROJECTS

Theia - Deep Learning Facial Login System *[Python (OpenCV, numpy, keras, scikit-learn, flask), PostgreSQL, AWS]*
- Developed a web application to automatically register attendance from video of people entering the premises
- Implemented face detection and face recognition incorporating one-shot learning with pre-trained FaceNet in Python achieving an average attendance accuracy of 85%

Article recommender system *[Python (pandas, numpy, scikit-learn, word2vec, TFIDF, flask)]*
- Developed a recommendation system from BBC news articles dataset in Python by computing article similarity using embedding vectors
- Deployed the system on AWS using Flask server

Time Series forecasting on Canadian Bankruptcy Data *[R, Time Series (SARIMAX, VAR, Holt-Winters)]*
- Forecasted Canadian bankruptcy rates with time series models (SARIMAX, VAR and Holt-Winters) in R
- Prepared a business report for a non-technical audience

Age detection from Facial Images *[Spark-ML, Python (pandas, numpy, scikit-learn), AWS (S3)]*
- Trained an ensemble of Random Forest and Logistic regression model using Spark MLlib in Python to detect age from face image
- Deployed the model using a distributed Spark ML pipeline on AWS EMR cluster

Twitter sentiment analysis *[Twitter API, Python (Vader Sentiment, pandas, numpy, flask)]*
- Performed sentiment analysis using Vader sentiment model in python on tweets scraped from Twitter APIs
- Built a web server using Flask to display 100 most recent tweets color-coded by sentiment score

SKILLS & CERTIFICATIONS

- **Languages:** Python, C++, Spark, R
- **Databases & Tools:** SQL(PostgreSQL, BigQuery), NoSQL(MongoDB), AWS(S3, EC2, EMR), Git, Tableau, Plotly
- **Techniques:** Logistic Regression, Linear Regression, Random Forest, Gradient Boosting, Neural Networks, Clustering, Collaborative Filtering, Time Series Models, NLP, A/B Testing
- **Certifications:** Deep Learning Specialization - Coursera