Vivekpandian Veerapandian

San Jose, CA (214) 802-4843 **vivekpandian08.github.io** vivekpandian08@gmail.com
github.com/vivekpandian08 linkedin.com/in/vivekpandian8 medium.com/@vivekpandian08

Data Scientist with **4 years of experience** in extracting customer and marketing insights from data points by building Linear and Non-Linear ML models. Established a $25k revenue E-commerce business by recommending data-driven solutions.

EDUCATION

The University of Texas at Dallas - ***M.S., Business Analytics (Data Science Specialization)***, 3.67 Jan 2019 – Dec 2020

- Mentored 15 Grad Students for ML Project competition, and **secured 3rd place**. Built end to end regression model, developed web app using FLASK that predicts house price and hosted the app in the cloud using AWS.

College of Engineering Guindy, India - ***B.E., Electronics and Communication*** Aug 2007 – May 2011

TECHNICAL SKILLS

Programming: Python (Scikit-learn, Pandas, Numpy, TensorFlow, PyTorch, Keras, Spacy, PySpark), R, SQL, SAS

Visualization: Tableau, Power BI, Shiny, Plotly, Streamlit

Databases & Bigdata: Snowflake, MySQL, PostgreSQL, Oracle, MongoDB (NoSQL), Hadoop, Hive, Spark

WORK EXPERIENCE

SuperWorld, United States – Augmented Reality Real Estate on Blockchain

Data Science Intern | R Sept 2020 – Present

- Developed an interactive web app in R-shiny that does **real-time end to end sentiment analysis** of SuperWorld's tweets

Ordermycake.in, India – E-Commerce B2C Platform to sell Bakery Products

Senior Data Scientist | Python, Tableau, AWS EC2, S3 Oct 2016 – June 2018

- Generated **18% increase in revenue to $5k** in 2017 by leveraging **NLP techniques** to analyze customers feedback
- Recommended optimized price for products by web scraping and data mining to analyze competitor product prices, leading to a **$3K increase in yearly revenue**
- Designed 3 A/B tests to identify the most engaging marketing campaign, resulted in a **30% increase in sales** and examined 8 layouts and designs, **increased Add to Cart rate to 40%**
- Led a 10-member cross-functional team to build an end to end B2C platform to expand the **operations from 2 to 8 cities** and presented reports to stakeholders using **Tableau** charts and dashboards

Data Scientist | R, Python Jan 2015 – Sep 2016

- Identified **4500 potential churn customers** by developing ML models using Logistic Regression, Decision Tree and **mitigated 36%** by offering them discounts
- Initiated a new payment method to solve delivery problems that satisfied customers and **increased unit sales to 6.5%**
- Revamped coupon mailing strategy for 3 customer segments by clustering using **K-means** and identifying the most engaging coupons leading to a **12% estimated increase in headcount**
- Built **machine learning pipelines using python**, optimized XGBoost model by adding new features to improve precision rate to 76% for the likelihood of purchase by analyzing consumer behavior

Cluster Wireless Software, India – Developed Software which facilitates M2M communications through IoT applications

Software Engineer **|** MySQL Mar 2012 – Sept 2014

- Designed **SQL queries** to extract information from IoT sensor data and identified anomalies by K-means clustering
- Analyzed product pain points and collaborated with a multi-functional team to develop robust solutions to meet client requirements, increased project conversion to 30%

PERSONAL PROJECTS (MACHINE LEARNING)

Can I predict your food? Maybe I Can! | Android SDK, Python (Deep Learning, GPU) Sept 2020 – Present

- Gathering food images using JavaScript Console Window and python, and labeling by "LabelImg" annotation tool
- Building a prototype that takes data from users and does **end to end ML to classify foods** using **TensorFlow Lite**

Uber Tweet Analysis| Python (Natural Language Processing, Unsupervised Learning) Mar 2020 – July 2020

- Data mined 5 years Uber tweets using Tweepy, stored in MongoDB, cleaned and processed data by establishing an **ETL pipeline,** and identified 3 key topics that customer mentioning using **LDA algorithm** and achieved coherence score of 0.36

Credit Card Fraud Detection| Python (Supervised Learning) May 2019 – July 2019

- Performed **quantitative analysis,** and **hypothesis testing.** Built an ML pipeline in PySpark to predict fraud transaction using SVM, KNN, Naïve Bayes Random Forest, and Neural Networks with **SMOTE** resampling and achieved AUC of 0.82