

# Prathamesh Satyawar Mahankal

Seattle, WA, USA | [psm1695@uw.edu](mailto:psm1695@uw.edu) | +1 (206) 581-5198 | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

## SKILLS AND ACHIEVEMENTS

- **Technical Skills:** Python (pandas, NumPy, SciPy, scikit-learn), R, SQL, Keras, Tensorflow, Spark, Hadoop, NoSQL, AWS
- **Tools and Frameworks:** Tableau, Power BI, Google Analytics, Advanced Excel, SSIS, Microsoft Office Suite
- Won the **Rising Star** award at **NASA's** International Space Apps Challenge Hackathon 2019 (out of 100+ participants).

## PROFESSIONAL EXPERIENCE

### Hi-Rez Studios, Seattle, USA

June 2020 – Present

#### Machine Learning Engineer Intern (Tensorflow, Pytorch, MySQL Workbench, Amazon Redshift, PySpark, PostgreSQL)

- Developed a data pipeline that uses complex feature engineering to generate a multi-purpose time-series based dataset for SMITE.
- Leveraged this dataset to build a Gradient Boosting Machine model that resulted in 4% better AUC in comparison to previous models.
- Built a sequence classification LSTM model and compared it with GBM, achieving a 12% and 8% rise in Precision and AUC respectively.
- Improved Item Recommendation System by A/B testing application of Collaborative Filtering & Content-Based Filtering techniques.

### Institute of Health Metrics and Evaluation (IHME), Seattle, USA

June 2020 – Present

#### Data Scientist (Python, SQL, MySQL, Linux, Putty, Git, Github, Agile)

- Building data cleaning and data preprocessing pipelines to enable researchers to model causes such as infertility and maternal health.
- Optimized data-streamlining processes that reduced redundancy by 10% and subsequently increased computation speed by 30%.

### The University of Washington, Seattle, USA

January 2020 - June 2020

#### Graduate Teaching Assistant (SQL Server Management Studio, Lucidchart, Tableau, Power BI)

- Helping students understand several DBMS concepts like Entity-Relationship Modeling, Database Design and Normalization, Querying, Transaction Management, Views, Stored Procedures, Functions, Joins, NoSQL, and Database Applications.

### Bank of America, Mumbai, India

June 2017 - August 2019

#### Data Analyst (Python, R, SQL, Advanced Excel, IBM Cognos, KDB+/Q, Tableau)

- Collected, cleaned, and analyzed massive unstructured equity trade data, as a part of a multicultural Big Data Analytics team.
- Demonstrated exploratory data analysis and visualization skills using Tableau, enabling stakeholders to make data-driven decisions.
- Spearheaded the development of multiple ETL pipelines for creating standardized data repository from disparate sources.
- Collaborated with cross-functional teams, contributing towards research, build, and deployment of various reporting components.

## EDUCATION

### The University of Washington, Seattle, USA

September 2019 – Expected May 2021

Master of Science - Information Management - Data Science

Relevant Courses: Natural Language Processing (NLP), Business Intelligence, Machine Learning, Statistical Modeling, Strategic Leadership

### Veermata Jijabai Technological Institute (VJTI), Mumbai, India

August 2013 – May 2017

Bachelor of Technology - Electronics Engineering

Relevant Courses: Computer Programming, Statistics, Applied Mathematics, Image Processing, Data Mining, and Business Intelligence.

## RELEVANT PROJECTS

### Image Retrieval Using Caption Generator (Microsoft Azure, Google Cloud Platform, Keras, InceptionV3, GloVe)

- Built an image retrieval mechanism on top of a custom image caption generator model trained on the MS COCO 118k dataset.
- Implemented transfer-learning using InceptionV3 model for image vectorization and pre-trained GloVe vector for word embedding.
- Evaluated and compared the performance of different model architectures using the BLEU score as a metric.

### Text to Speech Using Tesseract OCR Engine (Deep Learning, Neural Networks, Keras, Tensorflow, CNN, Tesseract OCR)

- Designed a system to capture an image, extracted all the text from it, and then converted this text file into a .wav file.
- Developed a script to binarize the image and pass it through a Tesseract OCR system using the pytesseract module.
- Incorporating additional features to recognize handwritten texts using Keras and Convolutional Neural Networks.

### Using Behavioral Science to Fight Climate Change (Clustering, Random Forest, Bagging, SVM, Classification)

- Developed a web platform that analyzes county-wise environmental metrics and uses inference-based Machine Learning algorithms to detect the exact problem faced by that county and suggest tasks to individuals based on the problem.
- Won the **Best Customer Validation** award among 100+ participants at the Techstars Startup Weekend Seattle hackathon.

### Customer Segmentation Analysis (kMeans, Trend Analysis, Unsupervised Learning)

- Implemented behavioral analytics techniques to understand customer trends and determine what drives customer loyalty.
- Segregated customers into segments using the Recency, Frequency, Monetary (RFM) segmentation model.
- Proposed personalized marketing strategies to tackle complex technical challenges and improve user engagement and retention.

### Text Classification of Political Quotes (NLP, Text Mining, Supervised Learning)

- Classified quotes using Naive Bayes model that gave 4% better accuracy (94%) than Logistic Regression and Random Forest ones.
- Analyzed the impact of performing tokenization using TfidfVectorizer instead of CountVectorizer generated unigrams and bigrams.