

Prathamesh Satyawar Mahankal

Seattle, WA, USA | 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, PySpark, 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

Microsoft, Redmond, USA

October 2020 – Present

Student Researcher – Capstone (*Graph Analytics, Azure Cosmos DB, Azure Machine Learning, Data Mining*)

- Building an end-to-end cloud-based platform that collects data from social media websites like Twitter, classifies it as real or fake using Azure Machine Learning, and then stores this tagged data in Azure Cosmos DB in graph format to facilitate graph analytics.
- Creating visualizations like network graphs to explain abnormalities in this data and thus analyze the spread of disinformation.

Hi-Rez Studios, Seattle, USA

June 2020 – Present

Machine Learning Engineer (*PyTorch, MySQL Workbench, Amazon Redshift, Hypothesis Testing, PostgreSQL*)

- Developed a batch processing data pipeline that applies complex feature engineering on data consisting of over 4 million users and generates a multi-purpose time-series based dataset for SMITE, our free-to-play MOBA video game.
- Collaborated with the Marketing team to leverage this data and build a Churn Prediction model that reduced attrition rate by 22%.
- Analyzed the player behavioral data and recommended a change in the location of daily deals which increased its sales by 35%.
- Measured the impact of increasing Battle Pass duration on player engagement and monetization using causal inference methods.
- Improved Item Recommendation System by A/B testing the impact of switching to item-based Collaborative Filtering technique.

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

June 2020 – September 2020

Data Scientist (*Python, MySQL, Linux, Git, Version Control, Agile, Performance Optimization*)

- Built data cleaning and preprocessing pipelines that enabled researchers to model non-fatal causes like Infertility and Hepatitis C.
- Optimized data-streamlining processes that reduced redundancy by 10% and subsequently increased computation speed by 30%.
- Utilized PySpark to distribute data processing tasks for maternal health pipeline and thus speed up the overall process by 40%.

Bank of America, Mumbai, India

June 2017 - August 2019

Data Scientist (*Python, R, SQL, Advanced Excel, IBM Cognos, Tableau, ETL, Data Warehousing*)

- Led the development of multiple ETL pipelines to create a standardized data repository aggregating data from disparate sources.
- Optimized the collection, cleaning, validation, and exploratory data analysis (EDA) of over 2 billion rows of daily equity trade data.
- Built the reporting infrastructure using SQL and Tableau to provide real-time insights into KPIs like Gain Pain Ratio and Sortino Ratio.
- Collaborated with cross-functional teams to design and build reporting frameworks that support financial products for equities.

EDUCATION

The University of Washington, Seattle, USA

September 2019 – Expected May 2021

Master of Science - Data Science

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

Veermata Jijabai Technological Institute (VJTI), Mumbai, India

August 2013 – May 2017

Bachelor of Technology - Electronics Engineering

Relevant Courses: Computer Programming, Statistics, Image Processing, Data Mining, and Big Data Analytics.

RELEVANT PROJECTS

Image Retrieval Using Caption Generator (*Deep Learning, Neural Networks, Microsoft Azure, Google Cloud Platform*)

- 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.
- Used the BLEU score metric to evaluate the performance of different Deep Learning model architectures and choose the best one.

Retail Business Process Improvement (*Power BI, Supply Chain, Project Management, Inventory Management*)

- Created and maintained user-centric dashboards that informed business partners on key metrics and performance measures such as Inventory Turnover Ratio, Average Transaction Value, Year Over Year Growth, and Sell-Through Rate.
- Identified and recommended improvements in supply chain process that boosted profits of the store by 30% over a quarter.

DGA Domain Detection System (*Python, Machine Learning, AWS Sagemaker, AWS Lambda*)

- Deployed a RESTful API using AWS API Gateway to send request to detect a DGA domain name from a FQDN name.
- Built an XGBoost Classifier with feature engineering on AWS Sagemaker and trained model on 15M rows dataset.
- Scaled the product such that it could produce 1 million predictions per minute using AWS Lambda and Sagemaker.

Using Business Intelligence to Improve eCommerce Business (*SSMS, SQL Server Integration Services (SSIS), Tableau*)

- Extracted actionable insights from broad, open-ended questions to influence product strategy and meet business goals.
- Developed and monitored Key Performance Indicators (KPIs) like Conversion Rate, Year-Over-Year (YOY) Growth Rate, Product Affinity, Customer Lifetime Value (CLV), and Customer Acquisition Cost (CAC) to drive business growth and performance.