

NAISARGI DAVE

• ndave@wpi.edu • www.linkedin.com/in/naisargidave

SUMMARY

Seeking a full-time opportunity as Software Engineer with Focus on Machine Learning, Big Data, Data Science and Data Visualization.

EDUCATION

Worcester Polytechnic Institute (WPI), Worcester, MA

MS in Data Science, GPA 3.88/4.0

Aug 2019 - May 2021

Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai, India

B.Tech in Electronics and Telecommunication Engineering, GPA 3.13/4

Aug 2013 - May 2017

Related Courses:

Machine Learning, Introduction to Data Science, Statistical Methods for Data Science, Data Visualization, Big Data Management, Information Retrieval and Social Web, Database Applications Development, Data Mining Business Applications.

SKILLS

Programming Languages: Python, R, SQL, HIVEQL, Scala, Java, HTML, JavaScript

Frameworks: PyTorch, MySQL, SQLite, SAP-HANA, HADOOP, MapReduce, HDFS, HIVE, PIG, Spark, MongoDB, Sklearn

Areas: Data Science, Data Analytics, Business Analytics, Data Visualization, Machine Learning, Statistics, Artificial Intelligence, Big Data, Database Management, Data Mining, Software Development

Tools: Qlik Sense, Tableau, Anaconda, SharePoint Online, IBM Modeler, Git, JIRA, MATLAB, VISUAL STUDIO, MS O365 Apps (Planner, Forms, PowerApps, Teams, Excel, Word, PowerPoint)

INDUSTRIAL EXPERIENCE

Data Visualization Co-Op, AbbVie, Worcester, MA, USA

May 2020 – Nov 2020

- **Change Over Application:** Developed an application using **SharePoint** and **PowerApps** with the Quality Assurance team for capturing the changeover details of equipment used in drug manufacturing processes.
- **Dashboards:**
 - **Talent Dashboard:** Represented AbbVie employee details and the **acquisition, attrition rates** using visualizations in a **Qlik Sense** Dashboard for the Strategic Operations team to monitor. Implemented data masking script using **Python** to handle sensitive data.
 - **Training Dashboard:** Created a dashboard for supervisors and employees across multiple departments to **track and highlight upcoming and past due requirements** to ensure compliance.

Developer, Reliance Industries Ltd., Mumbai, India

Aug 2017 – Jun 2019

CIO Dashboard:

- Created a dashboard for the **CIO of Reliance** to monitor performance of the teams.
- Integrated data from **flat files, SQL tables and SAP-HANA**, modelled it and used **Tableau** to visualize the Key Performance Indicators (KPIs).

Text Classification:

- **Automated the classification of user queries** entered in the Grievance Redressal Portal of Reliance by clustering and classifying them using natural language processing techniques, **Naive Bayes and Support Vector Machine classifiers**.
- **Improved the efficiency** of the team by approximately **30%** by automating the old manual process.

Team Ranking Scorecard:

- Developed a scorecard to track the performance of several teams and rank them using 17 KPIs.
- Performed **statistical modelling in Python** and developed **visualizations using Tableau**.

Procurement Spend Analysis:

- Performed **complex data modelling** for procurement spend statistics and analysis of the organization using **Hive on Hadoop - MapReduce** framework. The results of the analysis were visualized using **Zoomdata**.

ACADEMIC PROJECTS

Melanoma Classification, Kaggle

Jun 2020 – Aug 2020

- Detected **melanoma** among images of benign and malignant skin lesions from the SIIM ISIC Melanoma Challenge Dataset.
- Used **EfficientNet** for feature extraction, incorporated metadata, performed **data augmentation**, and image preprocessing to **crop out regions of interest** from the images. Achieved **80%** accuracy.

Human Protein Classification, Kaggle

Jun 2020 – Aug 2020

- Performed **multi-class classification** to identify all the types of proteins present in the cell images from Human Protein Classification dataset.
- Performed data pre-processing and data augmentation, used **Transfer Learning** with **pre-trained ResNet50** model to make predictions.

Car Review Analysis, Worcester Polytechnic Institute, USA

Jan 2020 – May 2020

- Created a **search engine** to fetch the most relevant reviews to the given user query using **BM25**.
- Performed **topic modelling using LDA**.
- Generated more accurate user ratings based on **Vader sentiment analysis** of the user reviews.

Twitter Data Analysis, Worcester Polytechnic Institute, USA

Jan 2020 – May 2020

- Performed **data mining of Twitter data** using twitter API and tweepy library for tweets containing the keywords “Donald Trump” and “Joe Biden”.
- Performed **exploratory and sentiment analysis** to figure out the general sentiment for each candidate and **predict the winner** of the 2020 presidential elections.

Travel Itinerary Application, Worcester Polytechnic Institute, USA

Jan 2020 – May 2020

- Developed an **Entity Relationship Model** and a **Database** using **SQLite** to capture the customer and travel details such as preferred transport, hotel, restaurant and tourist attraction.
- Created an **Android application** to allow user to **plan the trip and view their itinerary**.

PageRank, Worcester Polytechnic Institute, USA

Jan 2020 – Mar 2020

- **Crawled webpages** using BeautifulSoup, performed **text preprocessing using natural language processing** (NLP) techniques.
- Implemented **PageRank** algorithm to rank the webpages.

Skewed Join Optimization, Worcester Polytechnic Institute, USA

Aug 2019 – Dec 2019

- Used **Apache Spark and Scala** to optimize the join operation between two **large data sets** (13M and 0.1M records) with one of them **skewed**.

Human physical activity recognition model, Worcester Polytechnic Institute, USA

Aug 2019 – Dec 2019

- Performed **feature extraction** and developed a **human physical activity recognition model** for predicting the activity performed based on the person’s movement data
- Evaluated and compared several statistical learning methods including **Logistic Regression, Random Forest and Support Vector Machine**.

CERTIFICATION

Deep Learning Specialization, Coursera

- Completed the Deep Learning Specialization course taught by Professor Andrew Ng.
- Implemented Image Classification, Object Detection, Face Detection, Neural Style Transfer using Keras and TensorFlow frameworks.
- Worked on natural language data to build Neural Machine Translation, Trigger Word detection and Text generation models using Keras and TensorFlow.