

# ARPAN KORAT

[akorat@buffalo.edu](mailto:akorat@buffalo.edu) || (716)-400 5918 || [LinkedIn](#) || [GitHub](#) || [Portfolio](#)

## EDUCATION

### University At Buffalo, State University Of New York – Buffalo

Sept 2021-Present

MS – Engineering Science – Data Science

GPA: 3.92/4.0

**Courses:** Probability, Numerical Mathematic (Advance Linear Algebra), Data Driven Analyses, Statistical Mining

### Gujarat Technological University (GTU), Ahmedabad, India

Aug 2016-Aug 2020

Bachelor of Engineering – Information Technology

CGPA: 9.25/10.00

**Courses:** Big Data Analytics, AI, Data Mining and Business Intelligence, Data Compression and Data Retrieval, Numerical and Statistical Method for CS, Database Management Systems, Python Programming, Software Engineering

## TECHNICAL SKILLS

**OS:** Unix/Linux, Windows, MacOS

**Prog Lang.:** Python (Pandas, Matplotlib, Seaborn), R, Java, C++, SQL, MATLAB, JavaScript, CSS, C#

**Frameworks:** Flask, Bootstrap, ML (sklearn, TensorFlow, PyTorch, Keras, PySpark, MLflow, MLlib), Sparks, AWS

**Database:** MySQL, Oracle PL/SQL, RDBMS, NoSQL, PostgreSQL, Google BigQuery, HIVE, MapReduce

**Tools:** GitHub, PyCharm, PowerBI, Tableau, SQLite Studio, Google Data Studio, MS office

**Other:** Probability, NLP, Statistics, Data Visualization, Big Data, PCA, Hypothesis Testing, ML Optimization

## WORK EXPERIENCE

### Data Scientist and Business Analyst (Intern), The Spark Foundation

Apr 2021-May 2021

- Perform exploratory data analysis (EDA) on various domains (Sports, Retail) using Python and Tableau.
- Created storyboards and dashboards for various tasks such as spread of Covid-19 cases worldwide using timeline analysis.

### Data Analyst (Intern), Adis Technologies

Jun 2020-Nov 2020

- Worked on Microsoft SQL Server using T-SQL.
- Generated the performance matrix for the companies' production units.
- Check for the loopholes in the database: Redundancy of the data, outlier detection, etc. Make database system 8% more robust than the existing database.

## ENGINEERING PROJECTS

### Automobile Loan Default Prediction [\[Link\]](#) | Python, SQLite, Seaborn, Sklearn, Google-Colab, PCA

Oct 2021-Dec 2021

- Project aims to predict if the car loan borrower will repay the loan to the lender or not. (Domain - Finance)
- Normalized the raw data to 3-NF and built data pipeline using SQL. Used Seaborn library for EDA.
- Implemented a SMOTE analysis on data to balance the classes. Deployed ML models such as L1 L2 regularization.
- Performed Hyperparameter Tuning and optimization on model and increased the accuracy rate from 90% to 94.82%.

### Prediction of Term Deposit Enrollment [\[Link\]](#) | EDA, Python, Pandas, R, Statistical Analyses, ETL

Oct 2021-Dec 2021

- Aim is to enhance efficiency of marketing campaigns, organized by banks that sell Term Deposit subscriptions as product. (Domain - Finance)
- Analyzed large retail bank data. Performed Statistical concepts and generate statistical graphs to detect the abnormalities.
- Implemented various regression analysis and classification ML algorithms such as XGBoost, Random Forest, SVM, Logistic Regression, ANN, Bagging, gradient Boosting, to predict the potential term deposit holders. Accuracy – 87%.
- Reduced the false detection of potential term deposit holder by 7%.

### Objects' movement detection for Robot [\[Link\]](#) | MATLAB, Applied Probability

Nov 2021-Dec 2021

- Developed probabilistic algorithm to simulate the object and robot movements in the horizon. (Domain – Research)
- Calculated the expected time to catch the object via robot. Detected 3 type of object robot interaction.
- Designed MATLAB simulations to dynamically detect the object movement and robot will act accordingly.

## LEADERSHIP & ACHIEVEMENTS

- Won 3<sup>rd</sup> prize in 'App Inventors Challenge' in National level competition, sponsored by IEEE. (Technical Event)
- I was a Student Coordinator and Event Head in state level technical as well as in sports events during 2018, 2019.
- Team leader in project group during 2017, 2018, 2019.