

# ABHIMANYU GANGULA

Tampa, Florida | [gangulaabhimanyu@hotmail.com](mailto:gangulaabhimanyu@hotmail.com) | [LinkedIn](#) | [Tableau](#) | [GitHub](#)

---

## PROFESSIONAL SUMMARY

Aspiring Machine Learning Engineer with **2+ years'** experience as a Data Analyst. Well-versed with statistical data modeling and skilled at deploying appropriate ML models to deliver business value. Adept at translating actionable data insights into business outcomes. Proficient in **Python, SQL, Tableau, Machine Learning Algorithms, and libraries.**

---

## ACADEMIC EDUCATION

### University of South Florida

Master's in Business Analytics and Information System

Jan'21 - Dec'22

### University of Petroleum and Energy Studies

Bachelor's in Applied Petroleum Engineering with Specialization in Upstream

July'13 - May'17

---

## SKILLS

**Programming:** Python, R, C#, SQL

**Big Data:** Apache Hadoop, Apache Spark, MapReduce, Hive

**Statistical Analysis:** Hypothesis Testing, A/B Testing, Inferential and Descriptive Statistics, Exploratory Data Analysis

**Machine Learning:** Time-series forecasting, Neural Networks, Natural Language Processing (NLP), Predictive Modelling

**Libraries:** Pandas, NumPy, SciPy, Scikit-Learn, TensorFlow, PyTorch, Keras, Seaborn

**Visualization:** Tableau, Power BI, Excel, Visio (UML)

---

## WORK EXPERIENCE

### *Machine Learning Intern* – iQuest Solutions

Aug'22 – Present

#### Time Series Forecasting

- Built a multivariate **time series forecasting** model in **Pyspark** to predict the busiest airlines and recommend better flight routes for a given season using **ARIMA, SARIMA, VAR** techniques and **neural networks (RNN, LSTM)**
- Evaluated the models and working on deploying the model to production environment using **Docker**.

### *Data Science Intern* – FCCI Insurance

June'22-Aug'22

#### Predictive Modelling

- Developed **underwriting** and **claims models** by creating predictive features, utilizing external data, and applying statistical and machine learning techniques (**GLM, Random Forest, Extreme Gradient Boosting**).
- Devised pricing and underwriting strategies based on **predictive model scores**.

### *Data Analyst* – Tech Mahindra Ltd

May'17-Jul'19

- Mined, analyzed, and manipulated databases with 20 years of data using **SQL** and **Excel** for diverse business requirements to generate data reports and key business operations strategies.
  - Designed **interactive, data-driven dashboards** and **scorecards** using **Tableau** and business intelligence tools to monitor real-time data.
-

---

## RELEVANT PROJECTS

- **Sentiment Analysis of product reviews of an e-Commerce platform** May'22
- Built a classifier to predict the reviews of products from Flipkart (e-Commerce) into their respective classes.
  - Loaded the reviews onto a spark session and pre-processed the dataset.
  - Processed the reviews with *NLP* techniques like *Tokenization*, stop words removal and built a *tf-idf* vectorizer as input to machine learning models (Random Forest Classifier, Linear SVC and Logistic Regression).
  - Built the classification model and evaluated the models to dish out the best predictive algorithm.
- Tools & Technologies used:* Databricks, PySpark and Python
- **Influence of socioeconomic factors on Incarcerated population** Apr'22
- Built comprehensive data models such as Poisson distribution, MLE and Quasi-Poisson Distribution in *R studio* to analyze the influence of socio-economic factors on prison population.
  - Collected data from various sources such as *prisonpolicy.org*, *data.census.gov* and multiple websites and pre-processed data along with feature extraction to make the dataset ready for analysis.
  - Examined the correlation effects and skewness in the data with plots, graphs and data visualization techniques.
  - Presented the marginal effects of each variable and made actionable recommendations to mitigate the influences.
- Tools & Technologies used:* Tableau, R and R studio
- **EDA in Film Industry for a successful movie studio** Nov'21
- Webscraped data from websites such as *imdb*, *moviefone*, *boxofficemojo* etc., to extract data using *BeautifulSoup*.
  - Loaded data onto a dataframe using pandas and cleaned/pre-processed the data.
  - Visualized the data in Tableau to answer questions such as :
    - a. How much should a studio spend to make a successful movie
    - b. What genres and age groups should the studio focus on
    - c. What time of the year should a studio slate their releases
  - Made recommendations and presented results for running a successful movie studio.
- Tools & Technologies used:* Tableau, Python and Jupyter Notebook
- **Predicting Drill bit wear and Reservoir Formation using ML techniques** June'22
- Analyzed data (drilling & logging) provided by Equinor on Volve field (North Sea) made public in 2018.
  - Problem statement is two-fold: To predict drill bit wear (Regression) and Reservoir type (Classification).
  - Developed and fine-tuned the ML algorithms for Formation Classification using K-Neighbors Classification(12 ¼) and Gradient Boost Classifier(8 ½) with an accuracies over 92% and 76% respectively.
  - Compared different ML algorithms like Decision Tree, Random Forest and Ada Boost with accuracies over 95 % to predict drill bit wear with corresponding sections (26", 17 ½, 12 ¼, 8 ½) accordingly.
- Tools & Technologies used:* Python, Jupyter Notebook, Tableau, Pandas, Numpy and Seaborn
- **Predictive Data Model using Azure ML Studio** May'21
- Built a predictive model in Azure ML Studio for Bionique Inc. to develop wearable medical-grade devices.
  - Used Machine Learning algorithms such as Random Forest, SVM and AdaBoost to train the models.
  - Trained the model against previous datasets of 5 years and tested it against a quarter of the data to evaluate the performance of the model.
- Tools & Technologies used:* Azure ML studio