PAWAN KALE

M A C H I N E L E A R N I N G E N G I N E E R

Tata Technologies



SKILLS

# PYTHON LIBRARIES

OBJECTIVE

NumPy, Pandas, Scikit-learn, Seaborn, Matplotlib, Flask nltk, Gensim, TextBlob

# MACHINE LEARNING

Linear & Logistic Regression KNN, Decision Tree

Random Forest, Boosting Algorithm

SVM, Naive Bayes

K-Means Clustering, Hierarchical Clustering, PCA

# NLP

Ngram, WordCloud RAKE, YAKE

One Hot Encoding, Bag of Words, TF-IDF, Word2Vec

# DEEP LEARNING

ANN, CNN

Activation Function: Sigmoid, tanh, ReLU, ELU, Softmax

Optimization Function: BGD, SGD, MBSGD, AdaGrad,

AdaDelta, Adam.

SOFT SKILLS

Written and Verbal Communication Interpersonal Skills Teamwork, Collaboration Presentation Skills

With the experience of 3.6 years in the field of **Data Science**, now seeking for a challenging position to utilize my skills and abilities that offers Professional & Organizational growth. Possesses proficiency in **Python, Data Analytics, ML Modelling** and ability to manage complex tasks to deliver insights and implement action-oriented solutions to business problems.

Professional experience in Predictive Modeling and implementation of **statistical machine learning algorithms**.

PROFILE SUMMARY

Able to investigate **Data Visualization** and summarization conveying key findings.

Ability to make **EDA, data wrangling, Data Processing & Extraction, Data Mining using NLP(Parsing)**.

Knowledge in data management tools - **MySQL and MongoDB** databases.

Ability to use Web Scraping Tools such as

**Regular Expressions and BS4**. Knowledge of **AWS, Flask, Git & GitHub**

EXPERIENCE

# Machine Learning Engineer

***Tata Technologies, From- Nov. 2019 to Present***

# KEY RESPONSIBILITIES:

Understand, analyze and interpret the datasets. Develop, debug and maintain ML software application written in Python.

Design Machine Learning Models for different types of data and for different output types (Classification, Regression, Clustering).

Investigate the behavior of input and output data numerically.

Investigate and optimize model’s performance.

**PROJECTS**

**Bank Personal Loan System for Identifying Prospective Customers**

### Domain: Banking

The client wanted a classification model that would determine whether the customer will subscribe to the personal loan or not so as to make further campaigning of their product.

From the collected data; EDA, Data cleaning and preprocessing is carried out.

The insights were derived from preprocessed data to show which group was likely to avail Personal Loan.

Feature engineering, feature selection and parameter tweaking are explored with the objective of achieving superior predictive performance.

Use predictive model building like Logistic Regression, Decision Tree, Random Forest, SVM, AdaBoost, GradBoost, XGBoost to increase and optimize customer campaigning and there by revenue generation and other business outcomes.

Select a final model by evaluating the different models and deploy the model using AWS.

## Tourist Review Analysis using NLP

### Domain: Travel & Tourism

The client was renowned travelling and tourism Industry providing fixed and customize packages for tourist. There main concerned was low reviews rating on online-websites and less booking for fixed package.

Client wanted to know the insights of data generated from reviews and based on that a classification model for tourist reviews.

Explored the reviews and rating given by tourist.

Extensive Text pre-processing and cleaning is performed on the data by NLP techniques such as Tokenization, Normalization, removal of stop words, numbers and punctuations, lemmatization.

Positive and Negative word cloud was created. BOW was created with Count-vectorizer and TF-IDF.

Logistic Regression, Random Forest, Naive Bayes & SVM was trained and tested. Logistic Regression was finalized for its high accuracy.

## Prediction of the Casting Product's acceptance rate

### Domain: Manufacturing

The client was a leading company in manufacturing industry. Client wanted to predict percentage acceptance rate of production before manufacturing the products.

Client wanted to make decision based on model prediction. Last 5 years of data provided of similar components.

After collection of data EDA is carried out. Preprocessing and Feature engineering is done.

Different Regression Algorithm was tried to find best prediction. Random forest was selected based on Evaluation metrices.

EDUCATION

PERSONAL DETAILS

CONTACT

**M.Tech** | ESC | GCOEA,

Amravati | 2013 | 8.47 CGPA

**BE** | ENTC | SGB Amravati University | 2010 | 67% **HSC** | Maharashtra State Board | 2006 | 77%

**SSC** | Maharashtra State Board | 2004 | 79.20%

**Date of Birth:** 21/1/1989 **Languages:** English, Hindi, Marathi

**Achievements:**

1. Qualified GATE-2011 Exam With 89.57 Percentile
2. Two Times Represent Maharashtra as a Lead Player in National Archery

9172991683

[pawandkale@gmail.com](mailto:pawandkale@gmail.com) Ward No. 5, Gajanan Nagar, Nandgaon Khandeshwar, Dist. Amravati,

Maharashtra, 444708.