

SHUBHAM GHAG

MACHINE LEARNING ENGINEER

CONTACT

7620369720

ghagshubham31@gmail.com

WORK EXPERIENCE

ML ENGINEER

Digital Hackzone,
PUNE FEB-2022 -
TILL DATE (1.7
YEARS)

SKILLS

- Python:- Numpy, Pandas, Matplotlib, Seaborn, Scikitlearn, OOPs, Regular Expression.
- Languages: Python, SQL.
- Cloud Platforms: AWS
- Web stack: Flask.
- Operating Systems: Windows Database: SQL, MongoDB

ML AND DS

- **Machine learning**- Linear Regression, Ridge & Lasso, Logistic Regression, Naïve Bayes Classifier, KNN, SVM, Decision Tree, Random Forest, PCA, AdaBoost, Gradient Boosting, K-means Clustering, Hierarchical Clustering
- **NLP**- NLP -: Text representation techniques, understanding text, text classification, clustering skills, LLMs, genai.
- **Libraries**: nltk, spacy, gensim, textblob, langdetect.
- **Deep Learning** - Neural Networks, ANN, CNN, Transfer Learning, Back Propagation, Tensorflow 2.x, Keras
- **Techstat** : Bag of words, Tfidfvectorizer, word2vec, sent2vec, keyphrase extraction

CAREER OBJECTIVE:

Experienced and dedicated AI Engineer with a strong background in machine learning, data science. seeking a challenging position in a dynamic environment at the Data Science and AI Competency Center. I am passionate about building high-performing ML/AI applications, implementing innovative solutions, and optimizing processes to drive business success.

CAREER HIGHLIGHTS:

- A professional with experience in **Python, Data Science** and **Machine learning** with expertise in Financial and Healthcare domain projects.
- Able to investigate **Data Visualization** and summarization techniques conveying key findings **PowerBI**.
- Communicates findings and obstacles to team members achieve best approach.
- Ability to write a clean and production code with **Object Oriented Programming** in **Python**.
- Experience in Web Framework **Flask** Thorough understanding of Probability and **Statistics, Bayesian methods**.
- Experience in data management tools- Relational and **SQL** databases.
- Knowledge of Python's **Data Analysis** and ML Libraries. Implementation of Recommendation system using Collaborative and Content based filtering.
- Source code management and Version Control system using **Git** and **GitHub**.
- Strong **communication** and interpersonal skills. Ability to interact with customers with ease and professionalism.

PROJECT:

Project : Customer Segmentation and Churn Prediction for Insurance

Domain : Finance

RESPONSIBILITY :

- Develop, debug and maintain ML and AI software applications written in Python ecosystem, SciKit- Learn. Design ML and Deep AI Models for different types of data and different output types (Classification, Regression, Clustering)
- Create stream learning models for forecasting, prediction, classification etc.
- Investigate the behavior of input and output data numerically. Investigate and optimize model's performance.
- Solving complex problems with multi-layered data sets, as well as optimizing existing machine learning libraries and frameworks.
- Work within a multidisciplinary team to understand client's business, data, and requirements and develop the appropriate AI or ML solution within our proprietary AI Machine.

EDUCATION

B.E With 9.3 SGPA| Savitribai Phule
Pune University |

PERSONAL INFORMATION

- **Current Address:** Gulmohar Park Society, Survey No 131/2, Walhekarwadi, Chinchwad Pune 411033
- **Languages Known:** English, Hindi & Marathi
- **Marital Status:** Unmarried
- **Hobbies:** Listening music, Cricket

TECHNICAL SKILLS

- **DEEP LEARNING:** NEURAL NETWORKS, ANN, CNN, BACK PROPAGATION, TENSORFLOW 2.X, KERAS
- **MACHINE LEARNING:** Supervised Algorithms, unsupervised algorithm, Clustering, linear, logistic, Naive Bays, Decision tree, Random forest, EDA, Feature engineering, Ensembling, predictive modeling, Statistics, Visualisation, Data analysis, Data processing etc...
- **AWS:** EC2, Sagemaker, S3, Deployment
- **NLP:** Understanding, representation, classification & clustering of Text Libraries: nltk, spacy, gensim, textblob, langdetect, googletrans
Techstat: BOW, TFIDF,
- word2vec, doc2vec, sent2vec, keyphrase
- **Extraction Databases:** 3C(Command, Constrains, Clauses), CRUD operations, Subqueries, Joins **Maths &**
- **Stats:** Filter, Wrapper, Embedded Methods, P-Value, T-Test, Z-Test, ANNOVA test, Chi-Square Test, Info-Gain Test, Hypothesis Testing., statistics, probability, statistics, linear algebra, Gradient Descent

NLP PROJECT:

Project Name: Sentiment Analysis and Root Cause Analysis of Customer Reviews and Comments.

Domain: Hospitality and Food

RESPONSIBILITY:

- Analyze, study, and transform data science prototypes as required.
- Design Natural Programming Language applications. Review and choose the best and most appropriate datasets for the best-supervised learning methods.
- Use well-formed and effective text representation to change and adapt natural language into user-friendly features.
- Research and implement the correct algorithms and tools for natural language processing tasks.
- Create and develop NLP systems by using the requirements set out for the project.
- Ensure the developed model is properly trained and ensure substantive evaluation experiments are run. Undertake statistical analysis of the results produced and refine the model where required.
- Keep up to date in the rapidly changing field of machine learning by constantly researching and learning new techniques and information related to NLP.