



KAPIL MUDSHINGE

Kolhapur, India 416143 • +919595012501 • kapilmudshingej@gmail.com

EDUCATION

- **MSc (Statistics)** - 2015 – 2017

=====

Data Scientist with a strong statistical background and over **5+ years** of experience utilizing AI and ML algorithms to solve a wide variety of business problems. Proficient in selecting the appropriate statistical tools to address specific business challenges, such as predicting customer behavior and preferences (**churn prediction**), **risk analysis**, and **demand forecasting**. Skilled in building web applications that leverage persisted and trained models, utilizing Flask, Python, and deploying on GCP and AWS Cloud

ALGORITHMS

- **Sales Forecasting with ARIMA:** A project to forecast sales for pharma products the **Autoregressive Integrated Moving Average (ARIMA) model**.
-
- Linear Regression, Neural Networks, Logistic Regression, Random Forest, K-means clustering, PCA/T-SNE, Support Vector Machines.
-
- **Sentiment Analysis** on call audio using natural language processing (NLP).
-
- Deep NLP - LSTM, RNN, Auto Encoder Decoder, LLM
-
- Computer Vision Image processing using OpenCV Python.
-
- Build a real time pipeline for video stream using event base message system i.e Kafka, Cassandra.
-
- Object Detection Model using YOLO V2, V3, V4, Efficientnet Using data from real time video feeds, detect & report social distancing violations (Person detect, Person tracking, Heat map)

SKILLS

- > Data analysis and interpretation.
- > Statistical modeling and hypothesis testing.
- > Machine learning algorithms and techniques.
- > Time Series Anomaly Detection with LSTM
- > MLOps: Gitlab, CI/CD, MLflow, DVC
- > Programming languages: Python, SQL, R, Snowflake.
- > Web application development: Flask.
- > Cloud platforms: GCP, AWS, Snowflake
- > Docker, Kubernetes: Proficient in containerization with Docker and container orchestration with Kubernetes.
- > Kafka, Cassandra: Experienced in implementing real-time data streaming with Kafka and building scalable applications with Cassandra.
- > GCP/AWS: EC2, Kubernetes, Lambda, ECS, Glue, Containerization, CloudFormation.

EXPERIENCE

Senior Machine Learning Engineer, 09/2018 - Current

Quantiphi - Mumbai, India

Customer Segmentation Model :

- Developed a system that could effectively classify and group customers in the pharmaceutical industry based on various characteristics, enabling targeted marketing and personalized strategies.
- Conducted exploratory data analysis to gain insights into the data and identify relevant features for segmentation.
- Evaluated the quality and validity of the segments using appropriate metrics, such as silhouette score or within-cluster sum of squares.

AI & NLP Based Claim Processing Automation Solution:

- Builds Document classification model for sorting of claims.
- Document digitization pipeline for extract all the information from policy and apply text classification model and build NER entities classification model (BiLSTM + Pre Trained Embedding) related documents.
- Building **knowledge graph - Neptune DB**.
- Build inference pipeline and deployed in server less application (**Lambda, SQS**) NLP Based Solution.

Built a product which helps enterprise clients for test case analysis In TaSS:

- Real-time test cases comparison system aimed at improving the efficiency of storing and managing workflows (test cases) using NLP.
- Capture workflow events directly from the execution environment.
- Developed machine learning latent semantic indexing algorithms and models to analyze and compare different workflow configurations (Built a search engine which has capable give workflow comparisons in real time)
- To ensure scalability and performance, leveraged distributed computing and parallel processing techniques **Kubernetes (Docker, Unicorn-Server)**.

Football Analysis using Computer Vision:

- The objective of the project was to develop a system that could accurately track and analyze the number of people entering and exiting stores in real-time, providing valuable insights to store owners and managers.
- Developed techniques to detect and track individuals as they entered or exited the store premises.
- Used state-of-the-art deep learning models, such as **YOLO**, to accurately detect and track people (**SORT**).
- To obtain football statistics, the system maintained a count of the number of people entering and exiting the store in real-time. This information was then aggregated over time to generate insightful reports and visualizations.