

# D.Abhishakth

✉ abhishakth007@gmail.com ☎ +919550406928 🌐 LinkedIn 🔄 GitHub 📍 Guntur , Andhra Pradesh

## PROJECTS

---

### Symptoms Based Disease Prediction System 🔗

- **Description:** Developed a **Proof of Concept (PoC)** model for disease prediction based on user-reported symptoms. Performed hyperparameter tuning and optimized model performance using **GridSearchCV** across **41 disease classes**.
- **Challenges Faced:** Initially faced overfitting issues (**Training Accuracy: 100%, Test Accuracy: 70%**) and model bias due to class imbalance in disease distribution across training and testing sets.
- **Solutions Implemented:** Conducted a thorough program review and identified **data leakage**, where the target variable had been merged with the input features. Resolved this and introduced **Stratified K-Fold Cross-Validation** to maintain class distribution, improving the model's performance to **86% training accuracy** and **82% test accuracy**.

### Cloud-Based Sales Data Pipeline 🔗

- **Description:** Designed and deployed a cloud-based **ETL pipeline** that extracts Apple's regional product sales data from **S3**, transforms it using Python, and loads it into a PostgreSQL database hosted on an **EC2** instance.
- **Challenges Faced:** Encountered system-level pip restrictions, AWS CLI installation issues, and GitHub authentication failures during deployment.
- **Solutions Implemented:** Resolved package errors using virtual environments, configured secure **AWS IAM** credentials for boto3 access, and resolved Git push issues using personal access tokens and branch force-push.

### AI Powered Sensor Anomaly Detection & Reporting System 🔗

- **Description:** Developed a working prototype to monitor **51 industrial sensors** using semi-structured, **unlabeled** time-series data. Built model for anomaly detection, fault prediction, and automatic report generation.
- **Challenges Faced:** Model showed high accuracy for the majority class ("Working") but frequently misclassified faulty machines due to **class imbalance**.
- **Solutions Implemented:** Applied **SMOTE** to balance the dataset during training, leading to improved accuracy and overall model performance.

## INTERNSHIP EXPERIENCE

---

### Machine Learning Intern, *Slash Mark*

- **Description:** During my internship, I developed a machine learning-based celebrity image classification system to distinguish between images of Lionel Messi and Virat Kohli. Leveraging a range of classifiers, I was able to achieve a **training accuracy of 86%**, with an **82% test accuracy**, which highlighted the need for further optimization.
- **Challenges Faced:** Encountered challenges in model performance, including a significant discrepancy between training and testing accuracies, pointing to potential overfitting or insufficient data quality for generalization.
- **Solutions Implemented:** Worked closely with senior engineers to implement **GridSearchCV** for hyperparameter tuning and integrated multiple classifiers (SVM, Random Forest, Logistic Regression). Utilized **OpenCV** for image preprocessing and established robust **pipelines** for efficient model training and evaluation.
- **Outcome:** Successfully improved **training accuracy to 96%**, gaining valuable insight into model optimization techniques while collaborating in a professional team setting.

## EDUCATION

---

<b>B.Tech in Computer Science</b> , <i>Parul University, Vadodara, Gujarat</i> Current CGPA: 7.8	08/2021 – 04/2025
<b>Class XII</b> , <i>Sri Chaitanya Junior College, Repalle</i> Percentage: 91%	05/2021
<b>Class X</b> , <i>Jawahar Navodaya Vidyalaya, Maddirala</i> Percentage: 76%	05/2018

## SKILLS

---

### Technology & Tools

AWS EC2/S3, Git, Jupyter Notebook, Keras, Matplotlib, Pandas, PostgreSQL, Python, Scikit-learn, TensorFlow, GPU Based Neural Network Training.

### Libraries

Pandas, Numpy, Sklearn, Pickle, Keras, Matplotlib, Seaborn

### Fundamental Skills

Git, GitHub, Operating Systems, Computer Networks, Math (Statistics), DBMS, Basics of Cloud Computing (AWS), Cloud Architecture , EDA , feature engineering , model evaluation.

### Soft Skills

- Problem-Solving
- Collaboration
- Communication
- Adaptability

## CERTIFICATIONS/ACHIEVEMENTS

---

**Python Essentials-1, Python Essentials-2**, *Cisco*

**Machine Learning Engineer level-1**, *IBM*

**Excel - Mother of Business Intelligence**, *Code Basics*

**CS50's Introduction to Databases**, *Harvard University, Massachusetts*

**Data Analytics & Data Visualization**, *NPTEL*

**Machine Learning & Data Science**, *Corizo*