KHAIRNAR SAGAR

sagar.khairnarr786@gmail.com| LinkedIn | Github | Ph.: 7383052425 | Surat, Gujarat

PROFILE

Hands-on Data Scientist with 1+ year of experience in machine learning, statistical modeling, A/B testing, and data-driven experimentation. Skilled at translating messy data into actionable insights, improving product usability, and solving real-world business problems. Adept at working with cross-functional teams in fast-paced environments to identify growth opportunities, design experiments, and build scalable ML solutions.

EDUCATION

Uka Tarsadai University | Bardoli, India
B.Tech in Computer Engineering, **CGPA: 6.49** **(August 2019 - July 2022)**
Diploma in Computer Engineering, **CGPA: 7.50** **(June 2016 - April 2019)**

SKILLS SUMMARY

- **Languages:** Python, SQL, R, C#
- **Frameworks:** Pandas, Numpy, Matplotlib, Seaborn
- **Data & ML:** Scikit-learn, TensorFlow, XGBoost, Random Forest, Statsmodels
- **Analytics:** A/B Testing, Hypothesis Testing, Statistical Inference, Feature Engineering
- **Big Data:** Hadoop, Apache Spark, Google Cloud Platform
- **Tools:** Power BI, Tableau, Advanced Excel, MySQL, MongoDB, Jupyter Notebook
- **Soft Skills:** Analytical Thinking, Team Collaboration, Communication, Problem Solving, Data Storytelling

WORK EXPERIENCE

DATA SCIENTIST INTERN | RUBIXE -AI **June 2024 - January 2025**

- Built and deployed **ML models (supervised & unsupervised)** to extract **business insights** and improve feature performance.
- Designed and ran **A/B tests** to evaluate **new features** and **discount strategies**, directly informing product growth decisions.
- Modelled seasonal demand trends to forecast key organizational **KPIs** and reseller behavior patterns.
- Conducted **statistical hypothesis testing** to validate **model results** and support product changes with data.
- Identified supply chain bottlenecks and helped suppliers meet **SLA** targets using **real-time analytics**.
- Used **neural networks** to personalize recommendations, improving customer engagement and reseller success.
- Collaborated with **product managers** and **engineers** to roll out scalable data solutions and usability improvements.
- Processed large datasets using **Hadoop**, **Spark**, and **SQL** to derive meaningful insights at scale.

Key Achievements:

✔ Increased forecasting **accuracy by 20%** using time-series models
✔ Improved feature adoption by **15%** via **data-backed A/B testing** insights
✔ Contributed to building an **ML-driven** recommendation engine for personalized product offerings

SOFTWARE DEVELOPER | NANOMINE TECHNOLABS LLP **April 2022 - May 2024**

- Designed desktop applications using **C#** and **.NET Framework**, enhancing system efficiency by **40%**.
- Implemented **WPF application** with interactive **UI/UX designs**, enhancing user experience and accessibility.
- Developed a **billing system** with **MYSQL integration**, reducing manual effort in invoice generation by **90%**.

DATABASE ADMINISTRATOR | NANOMINE TECHNOLABS LLP **January 2022 - April 2024**

- Designed and deployed a **scalable web application** using **Laraval with MYSQL** for efficient data handling. Worked closely with front-end developers to integrate **dynamic UI elements**, enhancing overall user experience.

PROJECTS

CLEANED CLIMATE NASA SENTIMENT **March - April 2025**

- Developed a **machine learning model** to analyze climate change trends using **time-series data**.
- Processed and visualized **global temperature anomalies** and CO_2 emissions to uncover long-term patterns.
- Implemented **multiple regression techniques** to model relationships between **greenhouse gas levels** and **temperature rise**.
- Applied **data normalization**, **outlier handling**, and **feature engineering** to prepare clean, analysis-ready datasets.

- Conducted **correlation analysis** and **hypothesis testing** to validate climate-impact factors statistically.
- Trained and evaluated multiple models **(Logistic Regression, Random Forest, etc.),** achieving over **79% accuracy** in predictions.
- Created interactive plots using libraries like **Matplotlib** and **Seaborn** to highlight climate dynamics.
- Designed the notebook for **reproducibility** and **stakeholder communication** with clear markdown explanations and conclusions.

INSURANCE CLAIM PREDICTION **January - February 2025**

- Built a **supervised machine learning** model to predict the likelihood of insurance claim approvals using customer demographic and policy data.
- Analyzed **customer profiles** and **behavior** to identify key traits influencing claim acceptance, improving personalization for **B2C clients**.
- Delivered insights that can be adapted by **agents/resellers** to optimize targeting and reduce claim processing time.
- Performed detailed **exploratory data analysis (EDA)** to identify key factors influencing claim outcomes.
- **Engineered features** and **handled missing values** to improve model performance and data integrity.
- Trained and evaluated multiple models **(Logistic Regression, Random Forest, etc.),** achieving over **98% accuracy** in predictions.
- **Tuned hyperparameters** and **validated models** using cross-validation techniques to avoid overfitting.
- Used classification metrics **(precision, recall, F1-score, confusion matrix)** to assess model performance.
- Provided **actionable insights** that can help insurance providers reduce fraud risk and automate the approval pipeline.

PROTUGESE BANK DATASET **December 2024 - January 2025**

- Analyzed a **real-world marketing** dataset from a Portuguese bank to predict the success of telemarketing campaigns.
- Performed in-depth **exploratory data analysis (EDA)** to uncover customer behavior trends and campaign patterns.
- Cleaned and transformed data by **handling missing values**, **encoding categorical features**, and **balancing** the dataset.
- Built and compared multiple classification models **(Logistic Regression, Random Forest, XGBoost)**, achieving **95% accuracy**.
- Used **ROC-AUC**, **precision**, **recall**, and **F1-score** to evaluate model performance and select the optimal model.
- Provided data-driven insights to help marketing teams better target customers and optimize campaign strategies.

KRISHIEVA **April 2022 - July 2022**

- Developed an **AI-based recommendation system** suggesting fertilizers & pesticides based on **real-time weather data**. Integrated image analysis models to assess plant health and provide personalized insights.
- Built an **efficient backend** with PHP, Laravel, and MySQL, ensuring **seamless data processing**.

CERTIFICATES

CERTIFIED DATA SCIENTIST (IABAC) **April 2025**
CERTIFIED DATA SCIENTIST (NASSCOM) **April 2025**
ANALYZING AND VISUALIZING DATA WITH MICROSOFT POWER BI **March 2025**
DATA SCIENCE FOUNDATION (IABAC) **February 2025**
CERTIFIED DATA SCIENTIST (Rubixe -AI) **January 2025**
CERTIFIED DATA SCIENTIST (DataMites) **September 2025**
WORDPRESS **July 2024**