PRABAL KUMAR DEKA

Phone: (+91) 8638517995 | Email: prabal456.pkd@gmail.com | Guwahati, Assam| LinkedIn | GitHub

SUMMARY

Entry-level Data Scientist skilled in Machine Learning, Deep Learning, NLP, Generative AI, and LLM-based application development. Experienced in Python, SQL, model training, evaluation, data preprocessing, and building end-to-end ML pipelines. Adept at feature engineering, EDA, prompt engineering, and deploying models in cloud platforms.

PROFESSIONAL EXPERIENCE

Data & Research Analyst – Turing   **Sept 2024-Aug2025**
- Implemented end-to-end data preprocessing pipelines using Python, NumPy, and Pandas, processing 100K+ records and reducing data quality issues by 35% to support accurate model training.
- Conducted statistical analysis and hypothesis testing on complex datasets, delivering actionable insights that informed strategic business decision.
- Collaborated with cross-functional teams to translate business requirements into data-driven solutions and technical specifications.

Computer Science Faculty – Potential & Concept Educations LLP   **Mar 2023-Sept 2024**
- Designed and delivered comprehensive programming curricula for 100+ students in Class 11 & 12, covering Python fundamentals and database management (SQL queries, joins, views, CTEs).
- Conducted hands-on workshops on data manipulation and analysis using Python libraries, resulting in 90% student pass rate in board examinations.

PROJECTS

Cars 24x7: Car Damage Detection (Link)
- Built and compared **CNN, EfficientNetB0, and ResNet50** models to classify car damages into categories.
- Optimized **ResNet50** using **Optuna** for **hyperparameter tuning**, achieving **79% accuracy** on test.
- Developed a **Streamlit web app** for real-time image-based damage detection.

LinkedIn Post Generator – Generative AI Project (Link)
- Developed an end- to-end Generative AI application using **LangChain** framework and **Llama-3.3** model for automated content generation.
- Implemented prompt engineering techniques and fine-tuning to improve content relevance by 85%
- Built interactive web interface using **Streamlit** with real-time content generation capabilities.

Credit Risk Modelling – Prudent Partners (Link)
- Developed machine learning models (**Logistic Regression, Random Forest, XGBoost**) to assess credit default risk for loan applications and to assign credit score to each customer.
- Utilized feature engineering techniques and handled class imbalance using **SMOTE** and Ensemble.
- Implemented hyperparameter optimization using **Optuna**, improving model **AUC** score to **0.89**.

TECHNICAL SKILLS

- **Languages & Tools:** Python, SQL, MySQL, Jupyter Notebook, Pycharm, Google Colab, MS Excel, Power BI
- **Libraries & Frameworks:** Fast API, Pandas, NumPy, Matplotlib, Seaborn, NLTK, SpaCy
- **ML / AI:** Langchain, VectorDB, scikit-learn, PyTorch, Predictive Modelling, Statistical Analysis, NLP
- **DevOps & Cloud:** Streamlit, Docker, Git, GitHub, AWS Sagemaker

EDUCATION

Central Institute of Technology, Assam   **June 2018 - June 2022**
- B. Tech in Computer Science & Engineering   CGPA – 7.83

CERTIFICATIONS

- **Data Science & GenAI Bootcamp - Codebasics.io**