Kumlesh Kumar

Junior Machine learning Engineer

kumleshkumarofficial@gmail.com

9861838389

Berhampur, Ganjam

https://www.linkedin.com/in/kumlesh-kumar-9a0315338

github.com/kumlesh-18

EDUCATION

Computer Science and Engineering (Data Science)
National Institute of Science and Technology (https://www.nist.edu/)
August 2024 – August 2028 | Berhampur

XII (Science) – CBSE – 2022
St.Xavier's High School
April 2020 – May 2022 | Berhampur

LANGUAGES

Odia ●●●●●

Hindi ●●●●○

English ●●●○○

COURSES

Applied Machine Learning in Python
University of Michigan
April 2024 – June 2024
- Supervised & unsupervised learning algorithms
- Model training, testing, and evaluation techniques
- Feature engineering & data preprocessing

SKILLS

Programming & Core ●●●●○
- Python (NumPy, Pandas, Matplotlib, Seaborn)
- Object-Oriented Programming (OOP)
- Data Structures & Basic Algorithms
- SQL (data querying & analysis)
- Git & GitHub (version control)

Data Analysis & Preprocessing ●●●○○
- Data cleaning & handling missing values
- Feature engineering
- Data normalization & scaling
- Exploratory Data Analysis (EDA)
- Handling imbalanced datasets

Machine Learning ●●●○○
- Supervised Learning
  - Linear Regression, Logistic Regression
  - Decision Trees, Random Forest
  - KNN, SVM
- Unsupervised Learning
  - K-Means Clustering
  - PCA (dimensionality reduction)
- Model evaluation
  - Accuracy, Precision, Recall, F1-score
  - Cross-validation

Deep Learning (Junior level — realistic) ●●●○○
- Neural Networks (basic)
- CNN for image classification
- TensorFlow / PyTorch (basic usage)

PROJECTS

Customer Churn Prediction System
Classification & Business Analytics Project
April 2025 – May 2025
- Analyzed customer behavior data to identify churn patterns and business risks
- Preprocessed large datasets by handling missing values and categorical variables
- Built Logistic Regression and Random Forest models for churn prediction
- Optimized models using feature selection and hyperparameter tuning
- Improved churn prediction accuracy compared to baseline models

Tech Stack: Python, Pandas, Scikit-learn, EDA

House Price Prediction System
Machine Learning Regression Project
January 2025 – March 2025
- Built an end-to-end machine learning pipeline to predict house prices using real-world datasets

Machine Learning with Python
Coursera
January 2024 – March 2024
- Supervised & unsupervised learning algorithms
- Model training, testing, and evaluation techniques
- Feature engineering & data preprocessing

PROFESSIONAL EXPERIENCE

Data Science Trainee
Analytics & ML Training Program
August 2025 – January 2026
- Worked on hands-on machine learning projects including prediction and classification systems
- Conducted exploratory data analysis on structured datasets
- Built end-to-end ML pipelines from data collection to model evaluation
- Applied feature selection techniques to improve model accuracy
- Documented results and presented insights

Technologies: Python, SQL, Pandas, Machine Learning Algorithms

Junior Machine Learning Intern
AI Solutions Lab (Startup / Training Company)
February 2025 – May 2025
- Assisted in building and testing machine learning models for real-world datasets
- Performed data preprocessing including cleaning, normalization, and feature engineering
- Implemented regression and classification models using Scikit-learn
- Evaluated model performance using accuracy, RMSE, and cross-validation
- Collaborated with senior engineers to improve model efficiency

Technologies: Python, Pandas, NumPy, Scikit-learn, Jupyter Notebook

- Performed data cleaning, feature engineering, and exploratory data analysis to improve model performance
- Implemented Linear Regression and Random Forest models and compared results
- Achieved improved prediction accuracy through hyperparameter tuning and cross-validation
- Evaluated models using RMSE and performance metrics

Tech Stack: Python, Pandas, NumPy, Scikit-learn, Matplotlib

Handwritten Digit Recognition

Handwritten Digit Recognition
Deep Learning Computer Vision Project
June 2025 – July 2025
- Designed and trained a Convolutional Neural Network (CNN) for digit classification
- Used MNIST dataset for training and validation
- Achieved high classification accuracy through model optimization
- Visualized training performance and prediction results

Tech Stack: Python, TensorFlow/PyTorch, CNN, NumPy

PUBLICATIONS

Customer Churn Analysis Using Supervised Learning Algorithms
Data Science Case Study
May 2025
- Investigated customer behavior data to predict churn probabilities
- Applied Logistic Regression and ensemble methods
- Highlighted business impact of predictive analytics in customer retention

Predictive Modeling for Real Estate Price Estimation Using Machine Learning
Independent Technical Publication
March 2025
- Explored regression-based machine learning models for housing price prediction
- Compared Linear Regression, Random Forest, and feature-engineered approaches
- Demonstrated performance improvements using cross-validation techniques