Himanshi Maheshwari

Greater Noida, Uttar Pradesh - 201310, India
+91-9302165123 • maheshwarihimanshi18@gmail.com
himanshi-maheshwari-1690a625b/ • himanshimaheshwari

Experience

RYM Grenergy **Gurgaon, India**
Data Scientist *June 2024 - November 2024*

- Engineered an AI-powered hill irrigation optimizer leveraging Python, Scikit-learn, and optimization algorithms, enhancing water efficiency by 20% across varying elevations.
- Conducted real-time traffic data analysis with Apache Spark and Pandas on over a million data points across Gurgaon's road network to detect congestion patterns.
- Analyzed an AI-driven adaptive traffic signal timing system that reduced mean wait times.
- Deployed AI-driven irrigation pump selection models using Google Colab, Jupyter Notebooks, and Flask APIs, optimizing scalable real-time deployment across varying elevations and shortening deployment time by 15%.

Education

Bennett University **Greater Noida, India**
B.Tech in Computer Science Engineering *August 2022 - July 2026*
CGPA: 8.45/10.00

Projects

Scheme Research Tool:

Tools: Streamlit, Python, Pandas, NLP, API Integration

- Engineered a research tool for government schemes using NLP and real-time API integration, processing 100K+ scheme records with BeautifulSoup, Requests, and OpenAI's GPT models for precise text extraction and summarization.
- Directed an intuitive search interface leveraging TF-IDF and Named Entity Recognition (NER), achieving 95% accuracy in retrieving relevant schemes from 10K+ user queries.
- Built 15+ interactive visualizations using Matplotlib and Plotly, delivering deeper insights into scheme distributions, eligibility criteria, and funding patterns.

Customer Transaction Analysis:

Tools: Python, Pandas, Matplotlib, Seaborn, Power BI

- Conducted Exploratory Data Analysis (EDA) on 1M+ transactional records using Pandas, NumPy, and Seaborn, uncovering key customer behavior patterns and spending habits.
- Designed 20+ interactive visualizations with Matplotlib and Plotly, identifying purchasing trends, seasonal fluctuations, high-value customer segments, boosting business strategy insights by 35%.
- Compiled data-driven reports using Jupyter Notebooks and Power BI/Tableau, empowering stakeholders to make data-backed decisions, enhancing marketing efforts by 25%.
- Applied statistical analysis techniques, including hypothesis testing and correlation analysis, to assess customer preferences, leading to a 15% uplift in targeted sales campaign effectiveness.

Big Data Analysis of EV Sales:

Tools: Apache Spark, PySpark, Google Colab

- Processed and analyzed a 500K+ record electric vehicle (EV) sales dataset using Apache Spark (PySpark) transformations and actions, accelerating data processing efficiency by 40%.
- Executed data cleaning and preprocessing with Spark DataFrame operations, resolving 98% of inconsistencies, handling 10K+ missing values, and normalizing data to extract key sales trends and customer preferences.
- Implemented aggregations and group-by operations to examine regional EV adoption patterns across 50+ markets, detecting 5 major sales hotspots and distribution trends.
- Applied statistical analysis and machine learning techniques (e.g., K-means clustering and time-series forecasting) to predict EV sales growth with 90% accuracy, supporting data-driven strategic decision-making.

Publications

Transforming Truth: A Comparative Analysis of Transformer-Based Models for Fake Data Detection on the Web. *Conference: ICICCT-2024 (Published).*

Technical Skills

Programming Languages: C++, Python, HTML/CSS

Web Technologies: Flask, Streamlit, REST APIs

Database Systems: MySQL, MongoDB, SQL

Big Data & Cloud: Apache Spark (PySpark), Hadoop, Microsoft Azure, Cloudera

Data Science & ML: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, OpenAI GPT models

Natural Language Processing: TF-IDF, Named Entity Recognition (NER), BeautifulSoup, Requests, Hugging Face Transformers

AI & Optimization: Genetic Algorithms, Deep Learning, Attention Mechanisms, Hybrid AI Models

Data Visualization: Matplotlib, Seaborn, Plotly, Power BI, Tableau, Folium

Mathematical & Statistical Tools: Excel, Hypothesis Testing, Time-Series Forecasting, Correlation Analysis

Other Tools & Technologies: Jupyter Notebook, Google Colab, Tkinter, VS Code, Git, Apache Pig

Certifications

IBM Machine Learning (Certificate: Link): Completed an in-depth course on machine learning algorithms, covering supervised and unsupervised learning techniques.

Object-Oriented Data Structures in C++ (Certificate: Link): Explored object-oriented programming principles and data structures such as trees, graphs, and heaps in C++.

Exploratory Data Analysis (Certificate: Link): Learned data preprocessing, feature engineering, and visualization techniques for extracting insights from large datasets.

Big Data Analytics with PySpark (Certificate: Link) : Developed skills in processing large-scale datasets using Apache Spark and PySpark for distributed computing.