**Mahesh Singh**

[maheshsingh8214@gmail.com](mailto:maheshsingh8214@gmail.com) | [LinkedIn](https://www.linkedin.com/in/mahesh-singh-52989a271) | [GitHub](http://github.com/mahesh8214) | Ph: +918078661108

**Professional Summary**

Enthusiastic Data Scientist with a solid foundation in **Statistices**, **machine learning**, **data analysis**, and data visualization gained through impactful internships at Ineuron and Pwskills. Proficient in Python, SQL, and a variety of data science tools, with a proven track record of deriving actionable insights from complex datasets. Passionate about applying data-driven solutions to real-world challenges and contributing to Google's mission of making information universally accessible and useful.

**Technical Skills**

* **Programming Languages:** Python, SQL, Java
* **Machine Learning:** TensorFlow, PyTorch, Scikit-learn, Keras
* **Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn
* **Big Data Technologies:** Hadoop, Spark, Hive
* **Database Management:** MySQL, PostgreSQL, MongoDB, NoSQL
* **Cloud Platforms:** Google Cloud Platform (GCP), AWS, Azure
* **Tools:** Jupyter, Git, Docker

**Professional Experience**

**Data Science Intern | Ineuron**

June 2023 - August 2023

* Developed and optimized machine learning models predicting customer behavior, increasing prediction accuracy by 87%.
* Conducted comprehensive exploratory data analysis (EDA) on datasets exceeding 1 million records to identify trends and insights.
* Collaborated with senior data scientists to preprocess data, including data cleaning and feature engineering, enhancing model performance.
* Designed and presented data visualizations with Matplotlib and Seaborn, facilitating stakeholder understanding and decision-making.

**Data Science Intern | Pwskills**

January 2023 - May 2023

* Participated in a team project to analyze and visualize sales data, resulting in a 15% improvement in sales strategy.
* Developed Python scripts for data cleaning, transformation, and analysis.
* Utilized SQL to query and manage relational databases.
* Presented analysis results to the team using interactive dashboards in Tableau.

**Projects**

**Customer Segmentation using Clustering Techniques**

Tools: Python, Scikit-learn, K-means

* Implemented k-means clustering to segment customers based on purchasing behavior.
* Analyzed the segments to identify key customer groups and tailored marketing strategies accordingly, resulting in a 25% increase in sales.

**Sales Forecasting with Time Series Analysis**

Tools: Python, ARIMA, LSTM

* Used ARIMA and LSTM models to predict future sales trends for a retail company.
* Improved forecasting accuracy by 20%, aiding inventory management and reducing stockouts.

**Predictive Analytics for Healthcare**

Tools: Python, TensorFlow, Logistic Regression, Random Forest

* Developed a predictive model to forecast patient readmissions.
* Utilized various machine learning algorithms, including logistic regression and random forests.
* Achieved an accuracy of 85%, helping the hospital reduce readmission rates by 10%.

**Diamond Price Prediction**

Tools: Python, Scikit-learn, Linear Regression

* Built a model to predict diamond prices based on various attributes such as carat, cut, color, and clarity.
* Achieved a mean absolute error (MAE) of 0.4, enhancing pricing accuracy for a jewelry retailer.

**Diabetes Prediction**

Tools: Python, Scikit-learn, Logistic Regression, Random Forest

* Created a model to predict diabetes using patient health data.
* Achieved an accuracy of 78% and an F1 score of 0.75, aiding early diagnosis and treatment.

**Albanian Forest Fire Prediction**

Tools: Python, Scikit-learn, Random Forest, XGBoost

* Developed a model to predict the occurrence of forest fires based on meteorological data.
* Achieved an accuracy of 82%, helping local authorities to allocate resources more effectively.

**Wafer Fault Detection**

Tools: Python, Scikit-learn, SVM

* Implemented a machine learning model to detect faults in semiconductor wafers.
* Increased fault detection accuracy to 90%, reducing production errors and costs.

**Certifications**

* Google Data Analytics Professional Certificate
* IBM Data Science Professional Certificate
* Machine Learning by Pwskills

**Education**

**Data Science Master Course| Pwskills Online**

Completed: 26 February 2023

* Relevant Coursework: Machine Learning, Data Mining, Statistical Analysis, Big Data Analytics
* Academic Projects: Developed a predictive model for healthcare readmissions, achieving an 85% accuracy rate.

**Diploma in Computer Science**

Completed: 2024

* From: Goverment Polytechnic College, Kota, Rajasthan
* Relevant Coursework: Algorithms, Database Management, Programming Languages like c , python
* Academic Project: Student Attendance Menagement , sentitment analysis