|  |  |
| --- | --- |
| Phone: | +92 337 8031066 |
| Email: | [isaad1066@gmail.com](mailto:isaad1066@gmail.com) |
| LinkedIn: | [linkedin.com/in/saadkhalid123/](file:///D:\Saad's%20files\Cv\linkedin.com\in\saadkhalid123\) |
| GitHub: | [github.com/EngrSaad27](https://github.com/EngrSaad27) |
| Address: | Karachi, Pakistan |

# MUHAMMAD SAAD

## Data Scientist & AI Developer

### SKILLS:

### 

|  |  |
| --- | --- |
| **Data Analysis:** | **Excel**, **Power B**I, Power point, **My SQL**, Data Preprocessing, Matplotlib, **Seaborn**, Plotly, ggplot. |
| **Machine Learning:** | **Regression, Classification, Clustering algorithms.** |
| **Deep Learning:** | ANNs, RNNs, **CNNs**, **GRU**, and **LSTM** |
| **NLP:** | Types of phrases, Normalization, POS Tags, Text Preprocessing, Tokenization, BOW, **NER (Name Entity Recognition)**, TF\_IDF, **Embedding**, Text Classification, Sentiment Analysis. **RNN, LSTM, and GRU** |
| **Generative AI:** | **Transformers**, **LLMs**, Hugging Face, **RAG**, Graph RAG, RAPTOR, ROUGE, BLEU, Fine-tuning, LangChain |
| **App Development:** | **Streamlit, Gradio** |
| **Libraries:** | NumPy, Pandas, **Scikit Learn**, **Keras**, **Tensor Flow**, Stats models, SciPy, NLTK, Spacy, YOLO, **Hugging Face**, Pinecone |

### PROJECT:

**Customer Segmentation & Prediction:**

[Github Link](https://github.com/EngrSaad27/Customer-Segmentation-Analysis)

|  |
| --- |
| * Conducted exploratory data analysis**,** data preprocessing and feature engineering on a dataset with **4500 customer records** with **0.5M dataset rows**. * Implemented unsupervised learning algorithms, including K-means clustering**,** hierarchical clustering and DBSCAN clustering to segment customers into **4 distinct clusters.** * Built a model for customer behavior using clustering algorithm, achieving a **silhouette score** upto **80%**. * Created [**customer segmentation app**](https://customer-segmentation123.streamlit.app) by using streamlit to **reduce cost** of paid tools. |

**Optical Character Recognition (OCR) App:**

[Github Link](https://github.com/EngrSaad27/OCR-App)

|  |
| --- |
| * Used **end-to-end OCR model** which is trained with **580M parameters** consisting of a high compression encoder and a long-context decoder. * **GOT – OCR 2.0** model used in this app with **97%** precision and recall of **English** and **Chinese** language OCR on document level. * App is created on Gradio app tool by using this model which is trained on CPUs with **128 batch size** and **3 epochs.** |

**Optimal Pricing for Retail Products:**

Github Link

|  |
| --- |
| * **Developed a machine learning model** to predict optimal retail prices for over **500+ products**. * Applied **regression algorithms** such as Linear Regression and Random Forest and many more achieving a model accuracy of **98%** (R² score). |
|  | | |

**Employment Data Analysis Using MySQL:**

[Github Link](https://github.com/EngrSaad27/MySQL-Employement-Data)

|  |
| --- |
| * Performed Data Retrieving, Updating and Deleting, Mathematical, logical, stored procedure operations, Data limiting, grouping, ordering options. |

|  |
| --- |
| **EXPERIENCE:**  **North America Distributors | Jun 2023 – Nov 2024**  **Amazon Account & PPC Manager**   * Managed PPC of **2** **Amazon account** which was aimed to decrease ACOS. * Optimized PPC with team member and decreased ACOS upto **12%.** * Developed and maintained detailed **flat files** using Excel to streamline data entry processes for multiple product listings, ensuring accuracy and consistency across the board. * Conducted **in-depth analysis** of campaign performance metrics using **Excel**, evaluating factors such as conversion rate, **ACOS, and ROAS.** * Compiled and updated distributor price lists using Excel, providing a comprehensive and up-to-date reference for pricing and inventory management. |
| **Product** **Centrum | April 2022 – May 2023**  **Product** **Analyst**   * Analyzed wholesale products from every angle and found profitable products, generating **15 times** high in sales. * Managed team of **5 product researchers**, increased **4 times** more Productivity by evaluating more than **200 distributors’ products** and **100 branded products** daily using MS EXCEL. * Generated huge Sales in **2nd Quadrant**, resulting 15 times more sales which was up to **1500% increment** in sales. |

### Certifications:

|  |
| --- |
| **Standford** Certified Machine Learning Course. [Click here](https://www.coursera.org/account/accomplishments/verify/RER2HYRTB4JO) |

### Education:

|  |  |
| --- | --- |
| Master Diploma in **Data Science & AI** from **Boston Institute of Analytics** | **2024-2025** |
| **Bachelors in Chemical Engineering** from **University of Karachi** with 3.73 CGPA  ***Registered Engineer from Pakistan Engineering Council*** | **2019-2022** |