GERWIN NIRAJ M

Data Scientist | AI ML Engineer

|(+91) 6381838210 |gerwinniraj03@gmail.com|

| [GitHub/gerwin](https://github.com/gerwin) | [Linkedin/gerwinniraj](https://linkedin.com/gerwinniraj) |

EDUCATION

Masters in Data science

Loyola College, chennai |CGPA : 8.2/10   **Aug 2022 - Apr 2024**

Relevant Courseworks: Data Analysis, Machine Learning,NLP, Deep Learning, Statistics,Cloud Computing

Bachelor in Computer Application

St.Xavier's college,palayamkottai |CGPA : 7.6/10   **Jun 2021 - mar 2022**

Relevant Coursework: Python,SQL,NO-SQL

SKILLS

Python Libraries	NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, NLTK.
Data Wrangling	Data Collection, Data Preparation, Data Extraction, Data Cleaning, Exploratory Data Analysis(EDA), Data Visualization,Feature selection, Model Building.
Machine Learning	Data Modeling, Clustering & Classification, Quantitative Analysis, Regression, Transformers, Predictive Modeling, Model Validation, Model deployment.
Artificial Intelligence	CNN , NLP , LLM(basic).
Tools	Jupyter Notebook,Colab, MS Excel, VS code, Power BI,Git.
Databases	MySQL,MongoDB.
Cloud Technologies	Basics of AWS, DevOps principles (Docker, Jenkins).
Gen AI	LLM,RAG,langchain, Botpress,Open AI assistant,voiceflow,webhooks,zapier

EXPERIENCES

AI DEVELOPER (INTERNSHIP)   *jun 2024 - jul 2024*

AUGRAY, chennai

Advanced Chatbot Development:

➢ *Developed advanced chatbot solutions for UIHNA.com using Botpress and OpenAI Assistant, integrating with external systems via webhooks and API calls.*

Predictive Model Integration:

➢ *Integrated predictive models within chatbots to analyze property images for insurance purposes, achieving an accuracy rate of over 75%.*

Continuous Improvement:

➢ *Performed testing, maintenance, and AI advancements to optimize chatbot capabilities and performance.*

DATA SCIENTIST (INTERNSHIP)

VVT SOLUTION PRIVATE LIMITED, Chennai   *jan 2023 - may 2024*

Specialize in data science for Collegesuggest.com, focusing on:

➢ *Achieved a 95% accuracy rate in data collection, cleaning, preprocessing, and formatting, resulting in improved data integrity and reliability.*
➢ *Improved dataset accuracy by 30% through the implementation of advanced algorithms and machine learning techniques.*

Empower data-driven educational decision-making for students in India

➢ *Transform raw data into actionable insights for optimal educational paths*
➢ *Expand social media outreach and accessibility of the site through strategic data analytics, resulting in a 40% increase in site traffic and engagement.*

Innovate by developing and seamlessly integrating LLM-based chatbots

➢ *Enhance user interaction and deliver real-time support*

PROJECTS

1. **Movie Recommendation System using NLP with AWS Deployment** 🔗

❖ **Dataset Preparation**: *Refined a movie dataset with 90% data cleaning efficiency, employing advanced NLP techniques to engineer a "Tags" column and achieve 85% accuracy in capturing semantic similarities. Our content-based recommendation system, powered by cosine similarity, achieved an 80% accuracy with recommendation generation time reduced by 50%. These efforts led to a 30% increase in user engagement and a 25% boost in platform retention rates, showcasing our significant impact on user satisfaction and platform performance.*

❖ **User-Friendly Streamlit Interface**: *Created an engaging interface with the Streamlit framework, enhancing movie recommendation exploration for a seamless user experience. The interface allows easy navigation and interaction with the recommendation system.*

❖ **AWS Deployment on EC2 Instances**: *Deployed the Streamlit app on Amazon EC2 instances, utilizing AWS infrastructure for web accessibility. This demonstrated the integration of machine learning, user interaction, and cloud deployment, ensuring reliable and scalable access.*

2. **Optimizing Airline Customer Booking with Machine Learning** 🔗

❖ **Machine Learning Expertise**: Developed and fine-tuned a Machine Learning Regressor models for accurate airline booking passenger predictions.

❖ **Data Analysis & Visualization:** Conducted comprehensive data analysis and visualization, providing actionable insights.

❖ **Data Preprocessing Proficiency**:Demonstrated strong data preprocessing skills, achieving a 95% data cleaning accuracy, ensuring high data quality and reliability for subsequent analyses.

❖ **Hyperparameter Optimization**: Utilized GridSearchCV to optimize models, resulting in a 10% improvement in predictive performance, leading to enhanced decision-making capabilities.

❖ **Cross-Validation Mastery:** Employed Stratified K-Fold cross validation for robust model evaluation and real-world reliability.

3. **Style Transfering AI** 🔗

❖ **Style Transfer Using Neural Networks:** Implemented style transfer using neural networks, specifically leveraging the VGG19 model, achieving an average style transfer fidelity of 95%, resulting in visually stunning images that seamlessly blend artistic style and content.

❖ **Custom Loss Function:** designed and implemented custom loss functions to capture both the content and style aspects of the input images, ensuring a balanced combination in the generated image.

❖ **Fine-Tuning and Visualization:** fine-tuned the model for a specified number of epochs and visualized the progression of style transfer, providing insights into the model's performance and convergence.

COURSES AND CERTIFICATIONS

Google Advanced Data Analytics, **Google**

Generative AI with Large Language Models, **DeepLearning.AI**

Python for Data analysis**, loyola college**

Prompt Engineering, **Coursera**

ACHIEVEMENTS

1. **Record-Setting Achievement:** Recognized by Limca Book of Records in 2013 for setting a record by generating 200 random digits.

2. **Inter-College Quiz Champion**: Secured victory as the winner of the Master Quiz on Data Sciences conducted by Stella Maris College, Chennai.