

GUNPREET SINGH WALIA | DATA SCIENTIST

+91 8860963457 | itswalia20@gmail.com | linkedin/in/datawalia20 | github.com/Walia20



EXPERIENCE

Research Script

Aug, 2023 - Nov, 2023

Principal Data Scientist

Delhi, India

- Advanced skills in AI, NLP, Data Science and AWS led to a 30% system efficiency boost, saving clients \$14,000 and elevating satisfaction by 20%.
- Created Research Script, cutting plagiarism by 25%, adding \$5,000 value, and drawing in 35% more users through email marketing.
- Integration of AI/NLP raised revenue by 15%, about \$20,000, through premium subscriptions, enhancing investor appeal and market presence.

Satindra Metal & Graphics Private Limited

Aug, 2020 - Aug, 2023

Data Scientist

Delhi, India

- Drove a 36% surge in data distribution pipelines using AWS Lambda, S3, and AutoCAD API, bolstering operational efficiency.
- Enhanced next-year revenue by 25% via innovative tech integration: AWS, Deep Learning, NLP in Flask-based REST API.
- Revolutionized data analysis with Report Clustering using TensorFlow's Doc2Vec, BERT, and LDA in SpaCy/NLTK for profound insights.
- Streamlined deployments with Docker and Kubernetes on ECS Cluster, ensuring scalable and high-performance operations.

Tech Bae Pvt. Ltd.

Mar, 2020 - Aug, 2020

Data Scientist

Delhi, India

- Ideated and developed a new strategy to recommend Buyer's Choice items using Deep Learning in Automotive-search Platform Project.
- Led the technical architecture of NLP-based Chatbot and Car Recommendations which increased 28% User Engagement.
- Solved scalability problems by using r6g EC2-instances through Load Balancer, decreasing latency by 40% from 20sec to 5sec
- Created Data Pipelines using AWS GLUE and S3 connected through Lambda with Airflow for Batch Processes.
- Deployed Data Management system on EKS using Docker and Kubernetes.

Google Developers Student Club

Oct 2019 - Mar 2020

Cloud Tech Lead

Delhi, India

- Facilitated the Google Cloud Ready Program in the northern region of India, which successfully enrolled 250+ Students.
- As a lead organizer, I orchestrated monthly workshops and hackathons, fostering a community that grew by 50% in membership within a year

TECHNICAL SKILLS

Backend Tech: Python, Flask/Django/FastAPI, EC2, S3, Route53, Lambda, REST API

AI-ML : NLP, Data Science, Deep Learning, LLM, Generative AI, Transformers, Reinforcement Learning(RL)

Data Pipelines : SQL, DataBricks, PySpark, Redshift, GLUE, ETL, Redis, Airflow

Cloud Tech : Amazon Web Services(AWS), GCP, Azure

Devops : Git, CI/CD, Docker, Kubernetes, EKS, ECS

ML Platform : Tensorflow, Pytorch, NLTK, SpaCy, scikit-learn, HuggingFace

PROJECTS

DNA EDA App | EC2, Streamlit, S3, FastAPI, Pandas | [link](#)

- Takes a DNA Sequence Input for Nucleotide Composition of DNA
- Does EDA on the Input System and shows distribution of ATGC count
- Uses Pandas and Plotly for Data Analysis and Visualization

Prescription Drug Knowledge | EC2, OpenAI, NLP, Streamlit | [link](#)

- Takes input a Drug Name as described in Prescription
- Used Transformers based OpenAI API for Prediction
- Predicts Description, Doctor Confirmation, Consumption Steps, Dosage

Symptom Disorder Prediction | Facebook-bart, Transformer, NLP, EC2 | [link](#)

- Takes Inputs Symptoms of a Patient with few Disorder Options
- Predict the Probability of Disorder based on context of Symptoms
- Uses Zero-Classification for prediction of Disorder from FB-BART model

EDUCATION

● **Sharda University**
B.Tech

2019, Sept - 2023 Aug
Greater Noida, U.P, India

● **Gurusharan Convent School**
High School

2019, Mar
Delhi, India