Gurvinder Singh Yadav

Analyst/ML Engineer

🏠 Portfolio 🐙 Github in Linkdin ✉ gurvindernoida@gmail.com 📞 +919910517638

EXPERIENCE

DRIFT LAB INNOVATIONS | MACHINE LEARNING INTERN
July 2023 - Nov 2023

- ➔ Developed the a langchain powered sales assistant with Llama2 and streamlit UI.
- ➔ Created a Docker image with all the dependencies for the given models, Deployed them on Vast.ai instances.
- ➔ Did a basic training of Vall-E and GradTTS using libreTTS dataset
- ➔ Segmented and annotated speech data of multiple speakers for fine tuning using Whisper-AI

INDIAN KNOWLEDGE SYSTEMS | RESEARCH INTERN
Nov 2022 - Dec 2022

- ➔ Scraped data from world history & sacred texts website.
- ➔ Configured and ran a multinode spark cluster and ran pattern mining algorithms and found relationships between verses, prefixes and there meanings.
- ➔ Analysed the evolutionary pattern in these sacred text and tried to see how particular use of grammar, and vocabulary evolved over time.

PUBLICATIONS

RECOMMENDATION SYSTEMS | COLLABORATIVE FILTERING, MPIP, DASK DISTRIBUTED

- ➔ Designed & optimised MPIP similarity technique to run on a distributed architecture in an asychronous manner with low memory footprint.
- ➔ **Published Research Article Link**

PROJECTS

HATESPEECH ANALYSIS ON TWITTER USING APACHE SPARK | PYTHON, SCALA, APACHE SPARK, DOCKER
2022

- ➔ Applied Linear Regression, Logistic Regression, SVM, BERT Transformer models on twitter hatespeech dataset.
- ➔ Orchestrated an Apache Spark cluster using Docker and classified realtime twitter data into hateful and otherwise.

GRAPH ANALYSIS OF TWITTER | PYTHON, PANDAS, NETWORKX
2022

- ➔ Scraped, cleaned, preprocessed twitter tweets on the topics of major political leaders in India.
- ➔ Created a graph of with leaders as node and used centrality (degree, eigen vector, closeness, betweeness etc.) to analyse the influence of political leaders and their polarising nature by classifing them into hatespeech.

SKILLS

PROGRAMMING

Proficient:
Python

Experienced:
Bash • SQL • Scipy
SQL • Sklearn

Familiar:
JavaSript • Shell • C
Julia • Cpython • hive

LIBRARIES/FRAMEWORKS

Pytorch • Spark
Sklearn • Numpy • Pandas
Numba • Dask • Selenium
Hadoop • Matplotlib

TOOLS/PLATFORMS

Git • Github • Vast.ai • Docker
Google Cloud

EDUCATION

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY DHARWAD
BACHELOR'S DATA SCIENCE AND ARTIFICIAL INTELLIGENCE

Dec 2020 - Present | Dharwad, Karnataka
Data Science and Artificial Intelligence
Cum. GPA: 8.19 / 10

REFERENCES

Sagar Raikar,
Speech Scientist Uniphore
✉ raikarsagar007@gmail.com

Dr. B Bahadur,
Assistant Professor IIIT Ranchi
✉ bahadurbam43@gmail.com

Dr. Animesh Chturvedi,
Assitant Professor IIIT Dharwad
✉ animesh@iiitdwd.ac.in