



Tavleen Bajwa
📍 Bangalore, Karnataka 📞 +91- 9602870072
Bioinformatics Analyst @MedGenome
tavleenbajwa8@gmail.com
Portfolio: www.linkedin.com/in/tavleen-bajwa-04aa88198/

tavleenbajwa8
tavleen-bajwa-04aa88198
@tavleenbj

Skills: Python SQL Machine Learning (SciKitLearn) NLP (NLTK) Statistics Streamlit Docker FLASK RASA NLU

Experience

- MedGenome Labs Ltd** research.medgenome.com Jan 2023 to Present
Bioinformatics Analyst Bangalore, Karnataka
Contributing to the development innovative AutoACMG Machine Learning Software tool, aiding variant prediction as a part of Data Engineering team
Utilized GET & REST APIs to generate data resources for around 5000 genes (JSON & XML) for analysis followed by preprocessing, paving the way for accurate model predictive analysis and EDA (Seaborn, Matplotlib)
Spearheading development of an in-house GPT-based Causal Large Language model, leveraging transfer learning from a Hugging Face pre-trained BIOGPT model by on proprietary data for literature search automation
Designing and implementing "SalesBot," an AI-powered customer support management assistant and chatbot, aimed at enhancing user experience using RASA NLU framework
- iNeuron.ai** ineuron.ai Jan 2022 to Dec 2023
Data Science Project Intern Bangalore, Karanataka
MLOps, Python, ML(Regression), Heroku, GIT cli, VS Code
Led the end-to-end creation of a Machine Learning (Regression) FLASK Web app, implementing CI/CD pipelines and building Data Ingestion, Transformation, and Validation pipelines
Designed overall app architecture & constructed app-related packages
Conducted EDA, pre-processing and checked linearity, normality & autocorrelation on Housing Price Prediction data from UCI repository for a regression model, deployed as a web app on Heroku Cloud platform

Projects

- Crop Recommendation App**
🔗 github.com/tavleenbajwa8/crop-variety-prediction
HTML, Python, Procfile, ScikitLearn, NumPy, Pandas, FLASK, Pythonanywhere by Anaconda
- Aid Escalating Internet Coverage**
🔗 github.com/tavleenbajwa8/Aid-Escalating-Internet-Coverage-
Natural Language Processing model to foresee which web pages will receive extortionate, enduring traffic
Jupyter Notebook, Python, NLTK, Doc2vec, Word2vec, TF-IDF, SciKitLearn, Pandas
- Computation Drug Discovery using ChEMBL Database & Machine Learning**
🔗 github.com/tavleenbajwa8/ML-Drug-Discovery-Project-
Ensemble-based Machine Learning Model using ChEMBL database's bioactivity data of Acetylcholinesterase targets
Jupyter Notebook, Python, NumPy, SciKitLearn, Pandas

Education

- IIIT- Bangalore & IBAB** Jan 2022 to Jan 2023
PG Diploma, Big Data in Biology (GPA 8.82/10)
- Panjab University** Jul 2019 to Aug 2021
Master of Sciences (GPA 8.47/10)
- University of Delhi – RLA College** Jun 2016 to Jul 2019
Bachelor of Sciences (GPA 7.77/10)

Publications

- Microsoft's BioGPT: A GPT-based Language Model for Biomedical Text Processing**
Geek Culture Mar 2023
🌐 medium.com/geekculture/microsofts-biogpt-a-gpt-based-language-model-for-biomedical-text-processing-d490dbb5e232
- Examining the Intrigue of Generative Adversarial Networks (GANs)**
Geek Culture Mar 2023
🌐 medium.com/geekculture/examining-the-intrigue-of-generative-adversarial-networks-gans-c20953a4db7b
- Virtualization in Cloud Computing : Processor & Memory Virtualization (Part-1)**
Geek Culture Feb 2023
🌐 medium.com/geekculture/virtualization-in-cloud-computing-processor-memory-virtualization-part-1-8d1342720e51
- From Gradient Descent to Adam: Navigating the World of Optimization Algorithms for Deep Learning**
Geek Culture Feb 2023
🌐 medium.com/geekculture/from-gradient-descent-to-adam-navigating-the-world-of-optimization-algorithms-for-deep-learning-641d412fe6c7