Narayana Mayya
107, Passion Elite 1, Doddathoguru
Electronic city 1, Bangalore 560100
Projects - https://web-creations-for-ai-models.streamlit.app/

Phone: 944-928-0625
Email: narayana.mayya10@gmail.com
Current CTC: 7.5 lakh
Expected CTC: 12 lakh

Summary:

- Currently working as a data scientist in a startup company. Involved in every phase of software development.
- First rank in the ongoing sustainability hackathon
https://machinehack.com/hackathons/genpact_google_for_developer_sustainability_hackathon_lets_crack_the_climate_resilience_code/leaderboard
- Developed the in silico enthalpy of fusion prediction model. The performance of the model is better than any other enthalpy of fusion in silico prediction model in the industry.
- Proficient in ML and DL algorithms.
- Good understanding of data science life cycle.
- Utilized different data cleaning and data preprocessing methods.
- Multiple feature engineering techniques used and analyzed its effect on models.
- Deployed models in AWS instance using flask API.
- Have basic understanding of multiprocessing, multithreading, spark, streamlit, docker, git, web scraping, elastic search, data structure and algorithms etc and used them on some occasions.
- Experienced in hyperparameter tuning different models.

Work Experience: (https://www.linkedin.com/in/narayanamayya/)

Associate Data Scientist, Sravathi.ai, Bangalore   **Dec 2021 – present**
- Built a graph neural network based model for the prediction of solvation energy using pytorch framework.
- DPH model, Enthalpy of fusion model – Random forest, XgBoost, SVM, Linear regression, MLP.
- Solubility model – Optimization of hansen formula.
- Sigma profile prediction model built using embedding layer, LSTM and various fingerprinting techniques.
- Demonstrated the capabilities of various models by creating web app's using streamlit.
- Created modules for web app's and worked with web development team for its testing and deployment.
- Working on GPT-3 based model for predicting activity coefficient.
- Experienced in using chatgpt throughout the product development process.
- Used web scraping as well as other data extraction technique to get all drugs data from dailymed website.
- (slow server AWS free tier deployment) - http://15.206.167.42:8501/

Technical Support, Ooma, Fremont, CA (1 year, 5 months)   **Aug 2017 – Dec 2018**
Network Security Engineer, Celestix Networks, Fremont, CA(7 months)   **Jan 2017 – Jul 2017**
Systems Engineer, Hewlett Packard, Bangalore, India(2 years)   **Sep 2011 – Sep 2013**

Technical Skills:

- **Libraries** - Tensorflow, pandas, numpy, scipy, matplotlib, scikit-learn.
- **ML algorithms** - K-NN, Decision Tree, Logistic regression, Linear regression, Naïve Bayes, LinearSVM and KernelSVM, Random Forest, XGBoost.
- **DL algorithms** - MLP,CNN, LSTM, RNN, GRU,GAN, Transformer, BERT, encoder-decoder, Siamese network, Attention model, VGG, ResNet, Inception.
- **Others** – Python, Bagging, Boosting, Stacking, TuriCreate, SQL, Flask, Jupyter, Colab, GCP, log-loss, AUC, Accuracy, Confusion matrix, Precision, Recall, f1 score, multiclass and multilabel classification, word2vec, BoW, tfidf, matrix factorization, kmeans++, PCA, tsne, EDA, SQLite, feature selection, Jaccard distance, Page rank, Github, Unet, matplotlib.

Education:

Master's in Electrical Engineering, California State University Long Beach   **August 2016**
Bachelor's in Electronics & Communication Engineering, Sir.MVIT, VTU, India   **June 2011**

Narayana Mayya

107, Passion Elite 1, Doddathoguru

Electronic city 1, Bangalore 560100

Projects - https://web-creations-for-ai-models.streamlit.app/

Phone: 944-928-0625

Email: narayana.mayya10@gmail.com

Current CTC: 7.5 lakh

Expected CTC: 12 lakh

ML Projects:

- Quora question pair similarity

Advanced NLP features created using fuzzywuzzy library. Calibration performed to get better probability values for computing log-loss.

- Personalized cancer diagnosis

Tried response coding with modified Naïve Bayes on categorical and text data. Interpretation is important so performed feature selection to get top 100 features. Noticeable change with class balancing. Stacking model using mlxtend.

- Facebook friend recommendation

Mapping the graph data to supervised learning problem. Used networkx library extensively. Created features using concepts like pagerank, connected components, Adar index, Katz centrality, HITS score, SVD.

- Stackoverflow tag predictor

Micro-averaged F1 score for multi label classification. Binary relevance method implemented in scikit learn used here.

- Taxi demand prediction

Used kmean clustering technique to split the region. For each region time binning has to be done. Dask used to process large data instead of pandas. Tried fourier transform for time series data.

- Microsoft malware detection

Created basic features for .asm and .byte files. Running 200gb of data was a challenge. Used GCP for executing different models.

- Human activity recognition

Classical ML models performed better than LSTM because of less data availability.

- Self driving car (controlling the steering)

Implemented Nvidia paper along with addition of dropouts. Used openCV to create the demo output.

- Amazon fashion recommendation

Using the product title, brand and color I developed basic NLP based model as well as DL based model to recommend products

- Netflix movie recommendation

Implemented many techniques from the Yehuda Koren research paper. Used surprise library for implementing recommendation system algorithms like SVD, SVD++ etc.

Certifications:

- AI for Medicine Specialization - DeepLearning.AI
    - AI for Medical Diagnosis
    - AI for Medical Prognosis
    - AI for medical treatment
- Deep Learning Specialization - DeepLearning.AI
    - Neural Networks and Deep Learning
    - Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
    - Structuring Machine Learning Projects
    - Convolutional Neural Networks
    - Sequence Models
- DeepLearning.AI TensorFlow Developer Specialization
    - Convolutional Neural Networks in TensorFlow
    - Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning
    - Natural Language Processing in TensorFlow
    - Sequences, Time Series and Prediction
- Build Basic Generative Adversarial Networks (GANs) - DeepLearning.AI…**and many more**