VARSHINI REDDY

MACHINE LEARNING| NATURAL LANGUAGE PROCESSING | COMPUER VISION | DATA SCIENCE

CONTACT DETAILS

Dallas, TX
(469) 610 - 1610
varshini.thatiparthi@gmail.com

GitHub - varshinireddyt
LinkedIn - vrt4
Web - varshinireddyt

OBJECTIVE

Passionate Machine Learning Engineer with 2+ years of experience in software developing, data mining, data visualization, predictive modeling, machine learning etc.

KEY SKILLS

PROGRAMMING
Python, Java, R (Basics),C,C++.

DATABASE
SQL/ MySQL, Mongo DB, PostgreSQL.

BIGDATA
Scala, Pig, Map Reduce, Hive

EXPERIENCE

MACHINE LEARNING INTERN • MINDCURRENT • MAY 20 - PRESENT

- Worked on stress prediction model using Natural Language Processing (NLP), Machine Learning (ML), Deep Learning (DL) models.
- Using MySQL, NoSQL database, fetched the model information.
- Applied State of the Art Machine Learning Techniques like SVM, Logistic Regression, Random Forest, Extra Tree, Bayesian, XGBoost, Gradient Boosting Algorithms etc. in Scikit-learn for classifying the responses into business categories.
- Evaluated the ML Models using standard Metrics like Precision, Recall, F1 Score, Log Loss, Mean Square Error, and Root Mean Square Error.
- Designed, train, test, evaluated Neural Network architectures like Convolution Neural Network, Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Anomaly Detection, Sequence to Sequence (Seq2Seq), Transformer.
- Performed Exploratory Data Analysis using Jupyter Notebooks/ Pandas/ NumPy.

INSTRUCTOR • TRIO UPWARD BOUND • OCT 19 – PRESENT

- Taught Mathematics subjects (which includes Linear Algebra, Calculus, Statistics, Probability, Geometry, Trigonometry) to high school students and also helped them for SAT exam.

SOFTWARE DEVELOPER I • ACCENTURE • JAN 18 – DEC 18

- Involved in various phases of Software Development Life Cycle (SDLC) in the requirement gathering from the business team, architect, design, develop, test and deploy to production.
- Followed Agile Scrum process to streamline development process with iterative and incremental development.

DEPLOYMENT
AWS Sage maker, GCP, IBM Watson, Flask

PACKAGES
Scikit- Learn, NumPy, SciPy, NTLK, Beautiful Soap, Matplotlib, Seaborn, Jupyter Notebook, Tensor Flow, Keras, PyTorch, OpenCV, Scikit-Image, FastAI, Transformers, SpaCy.

VERSION CONTROL
Git

DATA VISUALIZATION
Tableau, Power BI, Excel

FRAMEWORKS/TOOLS
REST web services, Micro-services, Laravel, Django, Eclipse, Visual studio code, Intellij, Atom, Xcode, Jira, JSON

WEB TECHNOLOGIES
HTML, CSS, PHP, JSP, JavaScript

CERTIFICATIONS

LYNDA
- Python for Data Science

COURSERA
- Machine Learning with Python - IBM

EDUCATION

MASTER'S IN COMPUTER SCIENCE • AUG 19 – MAY 21 • UNIVERSITY OF TEXAS – ARLINGTON • GPA – 3.8
Member of NSBE, Women Who Code, Fab Lab.

BACHELOR'S IN COMPUTER SCIENCE • 2014 - 2018 • G NARAYANAMMA INSTITUTE OF TECHNOLOGY • GPA – 3.6
Member of Mac Lab(Launched by Tim Cook, Apple CEO), Rotaract.

PROJECTS

TEXT CLASSIFICATION/SENTIMENTL ANALYSIS.

TWITTER ANALYTICS FOR US AIRWAYS - LANGUAGE MODELING, ULMFIT
- Implemented Universal Language Model Fine Tuning (UMLFiT) using Fast AI.
- Performed Feature Engineering, Text Cleaning using Pandas, NTLK, and Keras.

KEYWORDS *ULMFiT, Transfer Learning, Fast AI, LSTM.*
LINK https://github.com/varshinireddyt/ULMFiT

TOXIC COMMENT CLASSIFICATION CHALLENGE - BERT, PYTORCH
- Conducted Data Cleaning, Text Mining, Vectorization to improve model performance.
- Tested multiple classification models such as a Random Forest, Logistic, Extra Trees Classifier, XGBoost, and performed hyperparameter tuning to optimize the model prediction.
- Achieved 97% accuracy with ROC, AUC.
- Implemented using PyTorch, performed fine tuning the pre-trained model to achieve better accuracy.

KEYWORDS *PyTorch, Bert, Transformers, Data Mining, Text Mining, Data Preprocessing, Vectorization, Random Forest, Logistic, Extra Trees Classifier, XGBoost, Bayesian, ROC, AUC.*
LINK https://www.kaggle.com/varshinithatiparthi/kernel64eed79506

PATENT ANALYTICS - UNSUPERVISED DATA. **TOPICMODELING**
- Performed unsupervised learning algorithms like multi - dimensional scaling, hierarchical clustering, and K - Means.
- Built Topic Modeling using Non - Negative Matrix Factorization (NMF), Latent Dirichlet Allocation (LDA), Latent Semantic Analysis/Indexing (LSA/LSI).

From Deep Learning AI

- Deep Learning Specializing
- Sequence Models
- Convolution Neural Networks
- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
- Neural Networks and Deep Learning
- Structuring Machine Learning Projects
- TensorFlow in Practice

DATACAMP

- Supervised Learning with scikit-learn
- Feature Engineering for Machine Learning in Python

REFERENCES

Available upon request.

KEYWORDS *Unsupervised Learning, Multi-Dimensional, Hierarchical Clustering, K-Means, NMF, LDA, LSA, LSI, Named Entity Recognition, Data Mining.*

STACK OVERFLOW. **TAG PREDICTION.**

- Suggested the tags based on the content, based on the question posted on the Stack Overflow.
- Calculated the accuracy using Precision and Recall.

KEYWORD *NTLK, Data Mining, Data Preprocessing, Pandas, NumPy, Scikit-Learn, Exploratory Data Analysis (EDA), Text Processing, Text Mining.*

CHATBOT. **SEQUENCE TO SQUENCE**

- Built a question - answering system using Keras, NTLK, Genism (Word2Vec) to do preprocessing of the data.

RECOMMENDATION SYSTEM

MOVIE RECOMMENDATION - NETFLIX DATASET.

- Implemented collaborative filtering techniques to predict movie ratings with Probabilistic Matrix Factorization (PMF) and nearest neighbor-based approaches using scikit-learn.

ANOMOLY DETECTION/ FORECASTING

MACHINE FAULT TOLERANCE PREDICTION - TIME SERIES, SENSOR

- Predicted fault tolerance of real time anomalies on univariate and multivariate time series data (sensor) using LSTM Auto encoders, K Means Clustering, SVM.
- Filtered out the noise by conducting Data Preprocessing and Feature Engineering.
- Pinpointed the times of fault and failure in the machine.

YOLO

CAR DETECTION.

- Applied YOLO to perform a car detection system using Tensor flow.
- Filtered with thresholds on class scores.
- Used non max suppression to get rid of anchor boxes with low scores.

CNN

FACIAL RECOGNITION

- Built a facial recognition using Convolution Neural Network with the Tensor flow.
- Performed Image Preprocessing, Augmentation.