VIGNESH THIAGARAJAN

7575 Frankford Road Apt. 3122 , Dallas | +1(469)571-3764 | vignesh@utdallas.edu | LinkedIn Profile |Github Link | Portfolio

EDUCATION

Master of Science in **Computer Science – Concentration: Data Science** *EXPECTED*
The University of Texas at Dallas – CGPA 3.78 *MAY 2021*
Coursework: Artificial Intelligence, Machine Learning, , Web Programming Languages, Design and Analysis of Algorithms, Big Data Management and Analytics, Database Design, Statistical Methods for Data Science

Bachelor of Technology in Electrical and Electronics Engineering *MAY 2018*
Amrita University – CGPA: 3.8
Coursework: C Programming, Object Oriented Programming, Embedded Systems, Digital Image Processing
Honors and Awards: Graduated with First Class, Distinction (Highest Honors)

TECHNICAL SKILLS

Programming Languages	:	R, Python , Java, C , Scala
Web Languages	:	HTML5, CSS, Bootstrap, RESTful API, XML, JavaScript, JQuery
Database Languages	:	SQL, PL/SQL, NoSQL (MongoDB)
Big Data Technologies	:	Hadoop, Apache Spark, Apache Hive, Apache Kafka, Impala, Amazon S3, Amazon EMR, Microsoft Azure Databricks, Azure Blob Storage
Tools/IDEs	:	Eclipse, IntelliJ, Visual Studio, Power BI, Matlab, LabVIEW

WORK EXPERIENCE

Data Engineering Intern, ANSYS Inc. *JAN 2020 – APRIL 2020*

- Automated Data Collection, Data Cleaning/ Preprocessing, Data Storage and Reporting through Pyspark and Azure Databricks.
- Utilized PySpark for processing of large amounts of data.
- Reduced run-time of most dashboards by 20%-50% and reduced run-time of main dashboard by 85% by optimizing queries.
- Utilized Power BI , Databricks and Jupyter notebooks to create visualizations.
- Created new workflows and documented all key changes.

Graduate Teaching Assistant, UT Dallas : *SEP 2019 – DEC 2019*

- TA for the Course CS 3340 at UT Dallas (Undergraduate Computer Architecture).
- Graded Assignments, projects and performed other student evaluations.
- Helped students with queries and providing additional materials and guidance during office hours.

ACADEMIC PROJECTS

Online Music Store Web Application (MongoDB , Express.js , AngularJS, Node.js)

- Developed Full stack Web application which implements an online music store.
- CRUD operations were implemented for shopping cart items.
- User registration, User login and logout, navigation bar, search by name and genre were implemented.
- User account details, admin details, music albums and other data were stored in MongoDB.

Text Summarization (Apache Spark, Scala, Amazon EMR, Amazon S3)

- Represented sentences in vectorized format and found the similarities between them using cosine similarity.
- Implemented PageRank algorithm to assign ranks to all sentences.
- Classified the sentences as either Summary text or non- Summary text using Decision trees.

Random Forest and KNN Classifier (Python)

- Implemented Random Forest Classifier from scratch and utilized to train a model on Kaggle's Breast Cancer dataset and achieved accuracies comparable to scikit Learn libraries for random forests (~98%).
- Implemented KNN classifier from scratch and obtained efficient classifier (accuracy ~95%) on the same dataset. Achieved results similar to those obtained with the scikit's inbuilt classifier.

CERTIFICATIONS AND ACHIEVEMENTS

- **Certified LabVIEW Associate Developer**- Received after undergoing **National Instruments'** training program.
- Graduated from Bachelor's degree with **highest honors (distinction)**, 2018.
- **Published** a technical paper in ***IEEE PEDES Conference, 2018***: Link to the publication in the IEEE website: https://ieeexplore.ieee.org/document/8707465