VAMSHI KRISHNA SHANAGONDA

(Willing to Relocate) NY | vamshi399@gmail.com | *+1 201-238-4414* | LinkedIn | GitHub | Files | HackerRank | LeetCode

EDUCATION

Pace University, **Seidenberg School of Computer Science and Information Systems**   **New York, NY**

MS, Computer Science and Engineering | **Concentration:** *Artificial Intelligence* | **GPA:** *3.83*   *May 2021*

Courses: Database Management Systems | Internet Computing | Data Mining | .NET | Artificial Intelligence | Algorithms & Computing Theory| Parallel & Distributed Computing | Bigdata Warehousing | Scalable Databases | Capstone research project

TECHNICAL SKILLS

Programming Languages:	R,C#,F#, Java,Python, Scala, C++, Typescript, Go, Dart, Kotlin, PowerShell, UNIX Shell, MATLAB, Lua
Database Management:	Oracle, MYSQL, SQL server, PostGresql, Redis, Cassandra, mongoDB, H2, Neo4J, SQLite, AWS Redshift, DynamoDB, Teradata, Snowflake.
Operating Systems:	Windows, Mac OS, Linux (Ubuntu, Red Hat), Android (OS + Apps), UNIX (Routers).
Platforms and Tools:	Eclipse, Visual Studio, VS Code, Qt, Android Studio, Pycharm, Colab, Firebase, JBoss, GlassFish, tomcat.
Cloud Technologies:	AWS, Azure, Google cloud platform.
Frameworks and libraries:	Angular9, ReactJS, Flutter, Cordova, jQuery, Meteor, Spring Boot, Hibernate, .NET, ASP.Net, Hadoop, Spark, NumPy, Pandas, Matplotlib, MLlib, Gensim, scikit-learn, H2O, PyTorch, TensorFlow, Theano, Caffe2, Keras, NLTK, spaCy, BERT, Django, Flask, Scrapy.

RELEVANT EXPERIENCE

Pace University | Graduate Research Assistant | STEM Institute   **New York, USA**

Email Spam Filter | (*OS, CSV, Math, Codecs, String, Pandas, Collections, Counter, sklearn, Python***)** | GitHub

- Designed using the various classifiers and models of machine learning using Python and selecting adversarial classifier approach to update Naïve Bayes based e-mail spam filters in response to attacks that try to evade the filter and to reduce spam.

24-puzzle | (*Qt toolkit, Win32, MFC, wxWindows, Windows GUI***)** | GitHub

- Designed a cross-platform application using Qt toolkit to find solution of a given 24-puzzle problem. Given initial state of a 24-puzzle problem, and the program finds the best solution, the one with the fewest steps.

Generic Genetic Algorithm | (*JavaScript***)** | GitHub

- Developed to solve both constrained and unconstrained optimization problems based on a natural selection process that mimics biological evolution. The algorithm which repeatedly modifies a population of individual solutions.

Genetic Algorithm for functional maximum | (*TensorFlow, Python, Numpy, Matplotlib, Machine Learning***)** | GitHub

- Developed a genetic algorithm for find out the maximum value of the given function $f(x)=x+10sin(5x) + 7cos(4x)$ in the range [0, 9].

Million Queens | (*C++***)** | GitHub

- Developed an algorithm to solve the N-queens problem (solve 3000000 Queen the time required for about 5s, particularly fast) using QS4 algorithm to solve the N-queens problem.

Human Facial Expression Recognition | (*Artificial Neural Network, Backpropagation, BPNN, facetrain, imagenet, hidtopgm, outtopgm, C, XV***)** | GitHub

- Developed to recognize the features of the face as well as facial expressions from database using Python Libraries. The name of the picture contains the userid, sunglass, pose, expression (anger, sorrow, happiness), and scale information of the picture. Using input face picture our algorithm works as 1) specific-person-recognizer, 2) sunglass-recognizer, 3) face-recognizer, 4) pose-recognizer, 5) expression-recognizer. Implemented Stepwise Linear Discriminant Analysis (SWLDA) for feature selection. SVM is used as the classifier for expression recognition.

Tic-Tac-Toe | (*Python***)** | GitHub

- Developed minimax and evaluation function which plays a perfect game. This Artificial Intelligence program will consider all possible scenarios and makes the most optimal move.

Lane-finding algorithm | (*Python, OpenCV, Numpy, CV2, glob, matplotlib, pickle, ipywidgets, moviepy, IPython***)** | GitHub

- Built using Camera calibration, distortion correction, perspective transform, Color/gradient threshold, image rectification, color transforms, and gradient thresholding. Identified lane curvature and vehicle displacement. Overcame environmental challenges such as shadows and pavement changes. Enhanced the pipeline with Keras using a state-of-art deep learning

architecture that is both extremely accurate and lean. Optimized the model overcoming environmental challenges such as shadows and pavement changes on video data from an automotive camera taken during highway driving.

Search engine for astronomical data | *Python, Scrapy, elastisearch*

- Built a specialty search engine for centralizing astronomical data. The search engine selectively crawls the web and ranks the web pages based on relevance. Compared the performance of the search engine with google custom search engine in terms of relevance and page ranking.

Document Classifier for Digital Research Papers | (*Python, Tensorflow, Keras, NLTK, Numpy, Pandas, sklearn, docker***) |** GitHub

- Built a model that classifies 10+ million WOS labelled research papers (scholarly big data) according to their subject area. The model which analyses the scientific documents using word embedding models like BERT and GloVe on TF-IDF ranked tokens. Evaluated performance using LR, SVM, RF, GRU, and LSTM.

Suggestion Mining from Online Reviews and Forums | (*Python, Keras, Matplotlib, NLTK, gensim***) |** GitHub

- Implemented recurrent neural networks (RNN's) to predict the semantic nature of the text from online reviews and forums. Evaluated performance using CNN, LSTM, and GRU and achieved an accuracy of 85%.

Project management tool | (*Python, Django, React, Redux, Express, npm, node.js, GraphQL, Docker, Kubernetes, Redis Server, sqlite3, tortoise ORM, channels, cryptography, daphne, pillow, Twisted, GitHub***) |** GitHub

- Developed an application in which all the clients, managers and teams can add projects/tasks, chat each other or in groups, track the progress of the projects/tasks, login/logout, settings page, file storage repository page for project/task related artifacts, FAQ, home/dashboard, calendar page with the added tasks pinned with all details, email functionality, image upload functionality, bot functionality. Session Management for effective planning, estimating and tracking team progress. Databases used are sqlite3, tortoise ORM, Redis server in docker.

Attendance automation tool | (*Python, MySQL, PyQt5 – GUI, HTML, CSS, KNN classification, Numpy, OpenCV***) |** GitHub

- Developed an application to use advance face detection technique by machine learning algorithm to automate whole process of managing attendance in educational institute. Application uses camera to capture one's image and compares it with the image stored in database followed by marking attendance. Instructor will receive an email with a list of students who were absent for a particular class and school admin can immediately view attendance details after each class. Lastly, our application reduces manual errors by consolidating student attendance which can be revisited when required making it easier for instructors to manage student attendance and even download the data in excel sheet. Used KNN algorithm in which training data is generated and then built the face classifier in which video stream data of a person is read using OpenCV and extracted face out of it and predicted with name and his respective UniversityID.

Breast Cancer prediction and detection |(*KNN,SVM algorithm, Numpy, sklearn, pandas, matplotlib, seaborn,sys,scipy***) |** GitHub

- Collected, Explored, Prepared the data. Trained a model on the data and evaluated, improved its performance using z-score standardization and testing alternative values of k.

Social Media Insight using Naive-Bayes and graph mining | (*python, spacy, NLTK, NLTKBOW, DictVectorizer, Naive Bayes classifier, sklearn, NetworkX, matplotlib***) |** GitHub

- Created a pipeline that takes a tweet mentioning "Python" and determine whether it is related to the programming language, based only on the content of that tweet. Used spaCy library for word extraction and analysis on natural language. Transformed the original text documents into a dictionary of counts using spaCy's word tokenization using NLTKBOW. Transformed these dictionaries into a vector matrix using the DictVectorizer transformer in scikit-learn which enables the Naive Bayes classifier to read the feature values extracted in the first step. Then, Trained the Naive Bayes (BernoulliNB) classifier and calculated the f-score using sklearn model_selection library using which we can accurately determine if a tweet using Python relates to the programing language nearly 70 percent of the time. With our initial set of users, friends/followers of each of these users is fetched. Our function takes a twitter user's ID value and returns their friends. graphs are made where some users are friends of other users (although not necessarily the other way around) using NetworkX and matplotlib library.

Data Analysis of Outlet sales data | (*python, pandas, numpy, sklearn, Regression(Linear, Ridge, Lasso, xgboost), metrics, r-score, Root mean square error, seaborn, pandas_profiling***) |** GitHub

- Used exploratory data analysis, confirmatory data analysis and predictive data analysis for mathematical and statistical Model building of outlet sales data. Analyzed the various factors affecting the target variable(outlet sales) with visualizations using python libraries which include model_selection, train_test_split, linear_model. Model Building: Predicting the values of sales for the given data using Regularization models. Applied Regularization Techniques on the data to check whether they are better than the basic linear model.

Hospital Management System | (*C#, ASP.NET, WinForms, SQL Server***) |** GitHub

- Developed a web application in ASP.NET. For patient: Patient Home, Current Appointment, Bills History, Treatment History, Take Appointment, Notifications, Feedback. Patient can request for only one appointment at a time. For Doctor: DoctorProfile, PendingAppointments, TodaysAppointmemts, HistoryUpdate, GenerateBill, PatientHistory. For Administrator: Admin Home, View Doctors, View Patients, View Other Staff, Search Box, Add/Remove.

Virtual Working Collaboration Tool | (*Flutter, Dart, Firebase, Firestore, Golang***) |** GitHub
- An application developed which can address the inconvenience of working in geographically different and distant locations, help the organization to develop projects efficiently and economically. Features include Login/Register/Logout, Group chats, Live Chats, Media sharing, Video Collaboration. Uploaded & Retrieved data stored in Cloud Firestore.

Apartment application | (*PHP, MYSQL, XAMPP***) |** GitHub
- Developed application where the user can add/delete his/her apartments. User can view the other listing of the apartments with images, price, owner contact and other details. Login/logout/home/settings/shotlisted apartments pages included.

Blockchain network | (*Hyperledger Fabric Composer, docker, Angular8***) |** GitHub
- In this we set up the Traders, TradingCards in a REST API server on the permissioned distributed Hyperledger Fabric framework. Developed an Angular4 application which uses the REST API to interface with this network.

Real Time Data Pipeline | (*Python, Scala, sqlite3***) |** GitHub
- Built Real Time Data Pipeline, with scalable storage and computation frameworks/technologies, using Apache Kafka, Apache Spark, PySpark on Apache Hadoop Cluster which is on top of Docker, PostgreSQL, Data Visualization using Django Web Framework and Flexmonster on Docker to generate insights out of this server data from real-time data centers which will be used by the server/data center monitoring people and they have to track these server's status regularly and find the resolution in case of issues occurring, for better server stability.

Website log Report | (*Scala, Java, Spark***) | Databricks |** GitHub

- Extracted critical information from Log file and generated vital reports from an ecommerce application to determine the factors that might be related to a successful eCommerce strategy. Generated session report which provides information about the session activity, referring to the actions that a user with a unique IP performs during a specified period. The number of user sessions determines the amount of traffic that websites receive. Generated pageview report, which determines how many pages were viewed during a specified time and which parts of your website attract the most attention. Generated new visitor report, indicating the number of new users that have visited the website during a given time and which keywords they used for searches. Generated domains report, target domains report, URL report, top IP addresses report, search query report. Also found out the mostly used Cellular Network technology, Mobile connection type, Payment type, Device screen resolution, Browser used for shopping, Device type etc.

World Development Indicators Analytics | (*Scala, Java, Spark***) | Databricks |** GitHub
- The World Development Indicators from the World Bank contain over a thousand annual indicators of economic development from hundreds of countries around the world. Generated reports of GINI Index, youth literacy rate, Trade as a percentage of GDP for China and India, Exports/Imports of goods and services, GDP per capita, Poverty alleviation, Life expectancy at birth, Urban population growth, Infant Mortality, 10 countries with lowest/highest average income, Average income from years-years in Rich/Poor countries, Average income in a specific year, Life expectancy in a specific country, G-7 country birth rates, World Per Capita income.

Infosys Technologies Limited   **Hyderabad, India**
Senior Systems Engineer   *September 2015 – December 2019*
- End to end ownership of deliveries to production – including analysis, planning, model, design, develop, code, test, debug, document and push to production. Facilitated Agile ceremonies (daily stand ups, Grooming, Sprint planning, Demo's, Retrospectives). Built and managed the Scrum Team project schedules, coordinated with Project managers, reported on Sprint statistics such as velocity, points taken in/accepted on time, burn down etc. Planned delivery of work; demonstrated active leadership within the team and agile practices. Hands-on lead for an agile team protecting then from outside distractions, removed obstacles, kept the team focused and on track, and assist where needed. worked in a continuous delivery model. Conduct design, code reviews, sign off code for push to production. Attended PI planning, manage risk, and issues log. Worked in JIRA & Confluence administration involving creation and management of custom, complex projects, workflows, process audits, security schemes, custom fields, dashboards, and reports configurations.
- Worked with Amazon EC2, Athena, Glue, EMR, Lambda, Aurora, Amazon S3, Amazon RDS, IAM, Amazon Elastic Load Balancing, Auto Scaling, Cloud Front, CloudWatch, Cloud Formation, Kinesis, Data Pipeline Services, AWS Batch Services, SNS, SES, SQS, VPC/SubNet/Security Group, ECS, CLI, Route 53 and other services of the AWS family. Transferred the data using Informatica tool from AWS S3 to AWS Redshift.

- Enhanced the existing product with newly features like User roles (Lead, Admin, Developer), ELB, Auto scaling, S3, Cloud Watch, Cloud Trail and RDS-Scheduling. Created monitors, alarms, and notifications for EC2 hosts using Cloud Watch, Cloud trail and SNS. Designed and configured Network Subnets, Route Tables, Association of Network ACLs to Subnets and Open VPN.
- Designed AWS Cloud Formation templates to create VPC, subnets, NAT to ensure successful deployment of Web applications and database templates. Created S3 buckets also managing policies for S3 buckets and Utilized S3 bucket and Glacier for storage and backup on AWS.
- Worked with Cognitive Services, Azure Machine Learning Service, Event Hubs, Power BI Embedded, Data Lake Analytics, Data Factory, HDInsight, Azure Analysis Services, Log Analytics Workspaces, Virtual Machine Scale Sets, Virtual Machines, Batch, Azure Functions, Disks, Virtual Machines (Classic), Container Instances, Container Registry, Kubernetes Services, Azure Cache for Redis, Azure Database for MySQL, Azure Database for PostgreSQL, Azure SQL Database, SQL Elastic Pools, Managed Databases, Azure Cosmos DB, Table Storage, Application Insights, Service Bus (Queues, Topics), Event Grid, API Management, Logic Apps, Relays, Azure Stream Analytics, Azure IoT Hub, Azure Maps, Digital Twins (Preview), Device Provisioning Services, Automation, Azure Backup, Media Services, Live and On-Demand Streaming, Notification Hubs, VPN Gateway, Public IP Addresses, Azure DNS, Virtual Networks, ExpressRoute, Application Gateway, Network Watcher, Virtual Network, Azure Firewall, Traffic Manager, Load Balancer, Network Interfaces, Connections, Key Vault, Storage Accounts, Azure Data Lake Storage, Storage Sync Services, Blob Storage, Queue Storage, Azure Files, Storage Accounts, Azure Search, Azure SignalR Service, App Services, WebJobs (triggered and continuous WebJobs), App Service Plans.
- Worked with Google cloud services which include Compute Engine, App Engine, Container Engine, Container Registry, Cloud Functions, Cloud Pub/Sub, Cloud Endpoints Frameworks for App Engine, Cloud Storage, Cloud SQL, Bigtable, Cloud Datastore, Cloud Spanner, Persistent Disk, BigQuery, Cloud Dataflow, Dataproc, Cloud Datalab, Google Genomics, Cloud Machine Learning, Cloud Vision API, Cloud Speech API, Natural Language API, Translate API, Google Cloud Virtual Network, Cloud Load Balancing, Cloud CDN, Google Cloud Interconnect, Cloud DNS, Google Cloud IAM, Cloud Resource Manager, Cloud Security Scanner, Stackdriver, Deployment Manager, Cloud Shell, Google Cloud Billing API.
- Worked on Spark streaming (Lambda Architecture), Kafka, Storm, Spark SQL, Tuning and Debugging the Spark Cluster (MESOS). Expertise on Hadoop, Kafka, HBase, Zookeeper, Hive, Pig, Splunk, Sqoop, Oozie, Storm, Flume, Cognos, Power BI, and other data modelling, reporting, analytics and visualization in Tableau and other ETL tools.
- Familiar with data architecture including data ingestion pipeline design, Hadoop information architecture, data modelling and data mining, machine learning, advanced data processing and data warehousing, data lakes, ETL workflows.
- Building and implementing architecture roadmaps for next generation Artificial Intelligence solutions around analytics and automation for our clients in different industries. Executing proof of concepts to assess the value of AI use cases.
- Architected modern Artificial Intelligence solutions using virtual agents, visual analytics, deep learning.
- Built a model, in python and in R, to classify the images of cats and dogs (Kaggle datasets) using Convolutional neural network (CNN) with a multi-channeled image input with accuracy of 85%. Here data augmentation is implemented to avoid overfitting of the images. Worked on CNN architectures which includes ILSVRC, LeNET, VGG16NET, GoogLeNet. Worked on Recurrent Neural Networks (RNN) to recognize patterns in sequences of data, such as text, genomes, handwriting, the spoken word, numerical times series data.
- Worked in modeling techniques which include OLS, logistic regression, GLM, Random Forest, Boosting, text mining, social network analysis, credit scoring, tree-based methods, time-series and forecasting, survival analysis, neural networks, nonlinear support vector machines, association rule learning, principle components analysis, K-means clustering, k-nearest neighbor classifiers.
- Data Mining, Modeling, Statistical Analysis, Business Intelligence: Collected, cleansed and provided modeling and analyses of structured and unstructured data used for major business initiatives.
- Worked in creating and using advanced machine learning algorithms and statistics: regression, simulation, scenario analysis, modeling, clustering, decision trees, neural networks.
- Employed machine learning and deep learning skills for multiple applications, including Computer Vision, 3D Reconstruction, 3D Modeling, Object Detection, Recommendation Systems. Worked on NLP models like LSTMs, Transformers (BERT, GPT-3). Solid understanding of ML techniques (parametric and non-parametric models, clustering, anomaly detection, evaluation metrics, interpretability methods).
- Worked on USB technology, analyzers and driver development on Windows and Linux. Worked on compiler architecture and optimizations. Root cause analysis and resolution of U-Boot, kernel, driver related issues.
- Familiarity and experience with HW interfaces and RTL design for ASICs or FPGAs.
- Worked with real-time OSs, multi-thread processing, and programming in ARM or similar processors.
- Worked with communication protocol implementations, such as Ethernet packet processing.
- Worked on TCP/IP, UDP, and network topology. Worked with socket programming.
- Worked on low-level system facilities: threading, I/O, signals, shared memory, VM, VFS, distributed computing including exchange topology, market gateways, and SORs.
- Proficient in software programming, artificial intelligence, machine learning, deep learning skills for multiple applications including Computer Vision, Recommendation Systems and Natural Language Processing.
- Worked with UNIX from kernel to shell and beyond, including system libraries, file systems, and client-server protocols.

- Developed an e-commerce web application on Spring Boot microservices and its tools includes kafka(producer, consumer), docker, Kubernetes, AWS lambda, ELK (ElasticSearch, Logstash, Kibana), Jenkins, Maven, Git, SonarQube, GitLab, Artifactory, Nexus repositories.
- Upgraded a web application to React(Redux, Hooks, Context, Auth0) that dealt with the health insurance, claims, financials, budget and its documentation and facilitated migration of data financials, payment processing services, record management services, and insurance industry services.
- Developed a hybrid application to enable a client to manage office space in IONIC3, ELECTRON, ANGULAR4, JavaScript, jQuery, JEE(RESTful Web Services + design patterns), Spring Boot and Maven using Xcode and tested the application in IOS, Android, Web, and Desktop client. Worked on Apache Camel, JMS, RabbitMQ, DevOps, Ansible.
- Developed a hybrid application named Ramplink+ in Xamarin (.Net framework) which involves scanning the bagtags and allocating them into the flight belly in the RAMP area. It's basically a weight and balance system which ensures the weight balance of the flight before take-off. Upgraded the WSDL and deployed the backend (JEE + spring + Hibernate) API in the IBM websphere