Sanath Narasimhan

LinkedIn:https://www.linkedin.com/in/sanath-narasimhan-016b2a135

Github: https://github.com/sanath-narasimhan

Website: https://sananara-aryabhata.netlify.com

Contact: +1 (682) 228-9743

Email: sanath.narasimhan@gmail.com

EDUCATION

Master of Computer Science (currently perusing),

University of Texas at Arlington - **3.667 GPA**

(Course work: Design and Analysis of Algorithms, Data Modeling and Analytical Techniques, Data Mining, Computer Graphics, Animation, Robotics, Neural Networks, Computer Vision, and Convex Optimization)

Bachelor of Engineering in Information Science (2013-2017)

Reva Institute of Technology and Management, Bangalore**,**

Affiliated to Visvesvaraya Technological University, Belgavi **– 69.34%**

SKILLS

Profeciency in Programming Languages		
Data Science/ Machine Learning	Open CV (python)	Intermediate
	R	Intermediate
	Matlab CV	Intermediate
	Python	Advanced
Web Development	HTML, CSS, PHP	Intermediate
	Python-flask	Intermediate
	JavaScript	Advanced
Application Development	Core Java	Intermediate
	C, C++	Advanced
	Shell Scripting	Advanced

- Familiar with machine learning and data mining approaches.
- Implemented various algorithms like linear regression, logistic regression, Decision trees, SVM classifier and Convolutional Neural Network classifier from scratch on practice datasets from kaggle and hacker rank.

Areas of interest:

Artificial General Intelligence, Human Computer Interaction, Natural Language processing, Image processing and Computer Vision.

AWARDS

· Secured all India rank 11 in the IEO 2016

· Won best use of Algolia API at the HackUTA event, October 5,6 – 2018

· Awarded the "Lone Star" Non-Resident Tuition Waiver scholarship for the Fall 2018, Spring 2019 and Summer 2019 terms By Department of Computer Science at The University of Texas At Arlington

· Nominated for Non-Resident tuition waiver Scholarship for the Fall 2019, Spring 2020 and Summer 2020 terms ByDepartment of Computer Science at The University of Texas At Arlington

CERTIFICATIONS

· Introduction to Machine Learning by Andrew ng on Coursera, September 2017

· The Python Mega Course on Udemy, November 2017

· Data Science Certification Training - R Programming by Simplilearn, January 2018

LANGUAGES

English (proficient), Tamil (Mother tongue), Telugu, Kannada, Hindi

Experience as Junior Data Scientist Intern at Enquero Global LLP

September 2017 - June 2018

Client name clustering

Dataset: Around 1,000,000 client names of Cisco. The names have a lot of variations and spelling errors.

A Machine Learning based algorithm to resolve a dataset of client names for an Elastic search engine. The dataset is subjected to unsupervised learning with a hybrid clustering (hierarchical clustering and k-means)with 57% accuracy. The goal was to have all names associated with a single client company would fall within the same cluster using Machine Learning approaches, but the data had too many spelling errors. To improve the accuracy of the clustering , natural language processing techniques need to be implemented.

SaaS Churn analytics

Dataset: Around 10,000 sample records mimicking the subscription status for Cisco Saas for 230 partners.

Data analysis on a dataset of Subscription data of various clients to predict the likelihood of churn. The dataset is sampled for previous churn trends and finding the key attributes by Principal Component Analysis to define churn causation. This is then applied on recent clients for analysis. The model was overfitting in anomaly detection with 55% precision on test set due to lack of data. The aim was to demonstrate proof of concept for a classifier that detect possible churn partners by calculating the likelihood of churn.

PROJECTS

Pill-em-all: (Masters Project for Data Mining course work)

Dataset: 53,767 unique reviews from multiple users of 2638 medicines for 709 conditions.

A medication review app developed in Python using flask for webapp development. The system is built on dataset available on Kaggle, a set of user reviews for medications with their symptoms with condition the medicine is usually prescribed for. The search engine was implemented using inverted index, giving an average retrieval time of 0.05 seconds. The classifier used is based on Naïve Bayes algorithm with hyper-parameter tuning of the smoothing factor, achieved on an average 60% accuracy across all classes. A content-based recommender system was implemented based on the given user ratings. Based on a user's search result, they are recommended medicines within the same category (condition associated with) that have greater ratings from other users.

Blink as security system: (Masters Project for Computer Vision course work)

Dataset: A minimum of 15 blinks at two different rates for each user is collected using OpenCV's face landmark detection function.

A new approach for user authentication through blinks developed in python. The approach involves creating two SVM classifiers for two different blink rates of each user, which lets them set a password using the trained blink rates. Each new user is asked to create a user ID and then blink a minimum of 15 times for each rate of blink they want to set. The blinks are detected using facial landmark detection algorithm available and some hard coded rules to ensure false positives don't occur while labeling frames due to the use of Eye Aspect Ratio (EAR), which is the ratio of the width of the eyes based on horizontal landmarks and the height from the vertical landmarks, at every frame.

The SVM classifiers take the Eye Aspect Ratio (EAR) in a window of 15 frames and uses the existing labels to learn the weights for detecting a unique blink rate. The average precision of the SVM classifiers for three unique users was 91.0367%.The password can have either of blinks in any combination and sequence, we use a hash table to map sequence of blinks to characters depending on the length of the whole password ensuring the encryption of the blink password while storing in the database.

Thesis Work:

My general research work involves studying current Deep-learning approaches through recent papers and implementing them. Our lab Human Data Interaction Lab run by Dr. Deokgun Park has one main goal, to develop an Artificial Intelligent system that uses the theory Hierarchical Temporal Memory (HTM) proposed by Jeff Hawkins. The first step involves creation of a language encoder and an AI that attempts to learn language similar to a human infant.

As part of my Master's Thesis I am integrating data from The pupil Labs Core Eye tracker to visualize gaze and fixation of the user on a 3D model in Unity3D environment.The environment also integrates data from Optitrack 13W Motion Capture system and Faceware Face motion capture device. The aim of the environment is to capture the natural interactions that occur between an infant and a mother in the process of learning language and exploring a small room with various interactive toys.