

SURESH SONWANE

@sonwanesuresh95@gmail.com

GitHub: github.com/sonwanesuresh95

Github page: sonwanesuresh95.github.io

LinkedIN: linkedin.com/in/suresh-sonwane-4ba712101/

EDUCATION

Bachelor of Engineering in Electronics & Telecommunication, University of Pune.

Aug'14 - May'19 Pune, India

Coursework: Machine Learning, Artificial Intelligence, Probability & Statistics, Algorithms and Data Structures, Differential Equations.

EXPERIENCE

✓ SIMPLIFAI.AI, Data Science Intern

Jul'19 - Aug'20 - Pune, India.

• EmailBot - Automated Email classification System

NLP, Python, Tensorflow • Automated email classification using deep learning project for Norwegian insurance firm.

• Wrote Flask APIs for training machine learning models for text classification and storing trained models, data, keywords in the codebase, wrote .yaml files. • Trained deep learning models for email classification and categorization in norwegian language. • Worked on optimization - wrote highly optimized python code by redesigning the algorithm, reduced inference timings from 15 min+ to 7 sec on api calls • Built batch pipelines for training and saving ml models for text classification purposes.

• Fretex - PDF Data Mining Project

NLP, Python, PyPDF2 • Worked on Fretex - A pdf data mining project for a Norwegian firm. Built a Rule based model for data mining from digital pdf Invoices/receipts using regex rule, multiple keyword matching, special keywords matching.

✓ SKNCOE, Teaching Assistant, ML, AI

Dec'19 - Mar'20 - Pune, India.

• Worked along with PhD Professor and assistant professor to teach machine learning foundations to final year students.

• Demonstrated techniques in Machine Learning e.g. data preprocessing, preparation, feature engineering, model training, model evaluation, inference in Jupyter Notebook.

✓ SKNCOE, Research Assistant, Machine Learning

Dec'18 - Mar'19 - Pune, India.

• Worked on a research project: Regression Benchmarking on real-world dataset using traditional learning algorithms.

• Benchmarked evaluation metrics and training times of multiple traditional machine learning algorithms including linear, non-linear, tree, ensemble, perceptron algorithms.

• Created Project Report on the same.

✓ Freelancing, Business Website

Dec'20 - Jan'21 - Pune, India.

• Developed a Full Stack website (information/bloggng) for a local business. Customized UI according to the clients orders, Used Python-Flask for backend, sqlite3 for database & html.

PoC: tulsibag-space-observatory.herokuapp.com

TECHNICAL SKILLS

- **Strongest Areas:** NLP, Machine Learning, Deep Learning, Data Structures, Algorithms, Software Design.
- **Languages (experienced in):** Python3
- **Languages (familiar with):** C, C++, SQL, HTML.
- **Tools & IDEs:** JupyterNB, Pycharm, linux shell, Visual Studio Code, Spark/EMR script, AWS notebook.
- **ML Libraries:** Sklearn, Tensorflow, Keras, Numpy, Pandas, matplotlib, Nvidia CUDA, OpenCV, scikit-learn, NLTK, spaCy, TextBlob, Gensim, Stanford CoreNLP.
- **Data Science:** Apache Spark (Pyspark), Pandas, Pymysql.
- **Deployment & Serving:** Flask, Docker, Heroku, AWS EC2.
- **Amazon Web Services(AWS):** AWS EC2, AWS EMR.

KEY PROJECTS

Elasticsearch Netflix Search Engine

Elasticsearch, Pandas, Python Nov'20 - Dec'20 Pune • Created a ranked search engine to search and retrieve records from netflix movies dataset. Ingested records into elasticsearch using python API. which searches for query in all fields and returns results in a descending order of match scores.

Information Retrieval System

NLP, Word2Vec, Gensim Aug'20 - Sept'20 Pune • Worked on building a NLP preprocessing pipeline+Word2Vec vectorization pipeline to find top 5 (in ~20) wikipedia passages semantically similar to a given search query

• Used cosine similarity to calculate similarity between search query vector and passage vector.

Contextual Question Answering System (org)

NLP, Deeppavlov, Python Jul'20 - Aug'20 Pune • Built a closed - domain contextual question answering system web application based on Deep pavlov, python and flask • Uses TF-IDF for information retrieval and R-NET for reading comprehension.

Sentiment Classification on imdb, pre-trained

word2vec NLP, Python, Word2vec, Gensim Mar'20 - Apr'20 Pune • Trained multiple deep learning models using keras embedding layer and pretrained word2vec model.

• Benchmarked performance metrics and training times on each training method.

Logging Visualization Tool, Application logs(org)

NLP, Matplotlib, Python, Flask Sep'19 - Oct'19 Pune • Built a framework that generates statistical data visualization based on application exceptions, using a regex based rule engine, by which the logs are categorized.

ACHIEVEMENTS

• 5 Star Python Programmer, HackerRank • Courses/certificates: Apache Spark(udacity), Andrew Ng (Stanford) Machine Learning course(coursera), Microsoft Edx Intro to Data Science(Edx), Deep Learning by Francis Chollet(Google) • **Research Paper Implementations, Tensorflow, Keras, python:** VGG, AlexNet and counting...(started Apr'20) • **Teaching and Research Assistant** at my undergrad college, 2019-20 • Top 15% of a kaggle competition (kaggle.com/sonwanesuresh95)