

SHREYAS BILIKERE SHANTHARAJU

+91-9916464703

shreyas.arjun.007@gmail.com

<https://linkedin.com/in/shreyas-bilikere-shantharaju-03671615b/>

Bangalore

Summary

A professional in IT with 10.3 years and about 5+ years of hands-on experience in Data science with delivering valuable artificial intelligence applications with machine learning, deep learning, computer vision, NLP, data mining, data extraction, deployment using python and 5 years of experience in python development, pyspark, ELK, SAP-Linux automation, and web & API automation testing.

Technical Skills

- ✓ **Data Science:** ML, NLP, ANN, RNN, CNN, Statistics
- ✓ **Dev-ops:** GitHub, Jenkins, Docker, Kubernetes, ELK, filebeat, Kafka, syslog, RabbitMQ, Logstash, Elasticsearch
- ✓ **Cloud:** AWS EC2, S3 Bucket, SNS, Lambda
- ✓ **Python Library:** NumPy, pandas, scikit-learn, SciPy, TensorFlow, Keras, Matplotlib, Seaborn, categoryencoder, NLTK, genism, feature-engine, creme, text blob, spacy, os, subprocess, requests, sys etc.
- ✓ **Web application framework:** Flask, Django
- ✓ **Operating system:** Linux (Centos, Ubuntu, Suse, red hat), Windows
- ✓ **IDE:** Visual studio code, PyCharm, Sublime
- ✓ **Databases:** MySQL, Couchbase

Project1: Outage prediction using word2vec and LSTM- Predictive maintenance (Cloud4c)

✓ **Problem Statement:**

Cloud4c Services Pvt Ltd offering public, private, multi cloud services. Cloud services facing down time due to various reasons. Our main aim is providing high availability without any down time. In order to achieve this, we need to predict future outages by using past logs data and prevent outages.

✓ **Business Impact:**

Preventing outages providing high SLA and avoiding paying penalty to customers for outages.

✓ **Description:**

- ✓ Requirement gathering and problem statement discussion with domain experts
- ✓ Data collection from RCA document and ssh remote servers through paramiko.
- ✓ Data analysis and visualization
- ✓ Data labeling which belongs to pre-outage and not outage data
- ✓ NLP data cleaning steps.
- ✓ Word embedding using word2vec
- ✓ Model building using LSTM
- ✓ Model performance evaluation
- ✓ Model hyperparameter tuning using Keras tuner
- ✓ Model deployment using GitHub, Jenkins, docker, AWS ECR, AWS EC2
- ✓ Model retraining approach using crème library in incremental learning

Project2: AI enabled – Document classification and text extraction (Cloud4c)

✓ **Problem Statement:**

Document digitization tool for insurance companies which does Document digitization (by reducing manual human work), Document classification and Entity extraction of Accord Forms, Loss run claims and Policy documents using Computer Vision, NLP and AWS Textract based on

customer needs.

✓ **Business Impact:**

This is reducing manual efforts and saving huge amount of time for insurance company.

✓ **Description:**

- Data collections such as PDF files and uploading into PDF to image converter.
- AWS textract used extract text from images and push into aws s3 bucket.
- Created **google pub/sub** for integrating each module TFIDF is used for text normalization.
- Classifying invoice or policy documents using random forest machine learning model
- Extracting the customer specific entities using Textract response and displaying it on UI.
- Deployed in AWS EC2 using docker, Jenkins pipelines.

Project3: Aspect based sentiment analysis classification (Wipro)

✓ Problem Statement:

Customer sentiment analysis is the automated process of discovering emotions in online communications to find out *how* customers feel about your product, brand, or service. It helps businesses gain insights and respond effectively to their customers.

✓ Business Impact:

Sentiment analysis is extremely important because it helps businesses quickly understand the overall opinions of their customers about the product and service. We can make faster and more accurate decisions.

✓ Description:

- Requirement gathering and architecture design of the project for implementation.
- Data Collection & storing: Data reading from DB2 database and placing in GCS bucket
- Data visualization & analyzing: Data visualizing and understanding using matplotlib and seaborn
- NLP Text Cleaning Steps:
 - Convert other language to English language
 - Tokenization
 - Lower case conversion
 - Remove punctuations
 - Spelling correction
 - Remove stop words
 - Remove all white spaces
 - Substituting multiple spaces into single space
 - Remove all single characters
 - Converting short cut words to full form
 - Remove URL's, HTML tags
 - Convert Emojis into meaningful text
 - Stemming and Lemmatization
- Feature Selection: Selecting important feature based on domain knowledge
- Aspect extraction using Topic modelling and cosine similarity.
- Model Creation: TD-IDF+ Ensemble techniques with 7 classifier ML algorithms and word2vec+LSTM
Model hyper parameter tuning using grid search CV and Keras Tuner
- Model Deployment using GCP, Pickle file, Flask, Jenkins, docker, Kubernetes, Google containerRegistry, Spinnaker Pipeline creation, auto build enables using webhook

Work Experience

Company Name	Designation	Period
Cloud4c Services Pvt Ltd	Senior Software Engineer	May 2020 – Till Now
Wipro Technologies	Senior Software Engineer	May 2018-May 2020
CenturyLink India Pvt Ltd	Systems Engineer	January 2016 – April 2018

TATA Consultancy Services Ltd	Systems Engineer	December 2012 – December 2015
-------------------------------	------------------	-------------------------------

Certifications

✓ AWS cloud practitioner

Education Summary

College	Qualification	Subject	Percentage/Grade
BIT, Bangalore (VTU University), Karnataka	Bachelor Of Engineering (B.E)	Electronics & Communication (ECE)	67.23%