Yashaswini P

+91-9480747213 |yashaswinichelur@gmail.com |
https://www.linkedin.com/in/yashaswini-p-b835ba1b5/

SUMMARY

A curiosity-driven data scientist, eager to leverage machine learning and data analytics to extract meaningful insights, make informed decisions and solve challenging business problems. I ensure to contribute with my knowledge, logical thinking and analytical skills towards the consistent growth and development of the organization, and enhance my experience through continuous learning and teamwork.

EXPERIENCE

Intern –Innodatatics   Mar2020-May2020

- Worked in a team of 5, developed algorithm to retrieve top 5 questions and answers based on the user given keyword.

- Dataset was given in Json format, priorly, change the dataset from json to csv format.

- An interface to which takes text as input, performed Data cleaning, Exploratory Data Analysis, Data visualization, finding tuples, tfidf matrix, vectorization and cosine similarity score.Using the similarity scores print the top 5 questions and answers.

- Deploy the model on 'heroku' platform.

PROJECTS

Car's MPG Prediction   Jan2020
- Dataset consists of 81 records and 5 variables.
- Performed Exploratory Data Analysis, Data visualization and developed Random Forest Regression model to predict the MPG of the car with accuracy 95 per cent.
- Github Link-https://github.com/yashu181/Car-s-MPG-prediction.git

Statistical Survey of Production, Distribution and Utilization Pattern of Bio-Inputs
 (M.Sc. (Agri.) Thesis from Jan2019 to July2019)
- The study was attempted to analyze the growth in production using secondary data, distribution of bio-inputs over different seasons and impact on yield due to bio-inputs usage by considering primary data.
- Production of Bio-inputs data was collected from 1999 to 2018. Primary data was collected by using pre tested schedule.
- Multistage purposive sampling method was followed.
- Statistical tools used are Compound Annual Growth Rate (CAGR), Coppock's Instability Index (CII), Correlation, t-test, Paired t test and chi-square analysis.

SKILLS

- **Statistics** –Hypothesis testing, estimation, probability theory, time-series analysis, statistical modeling.
- **Machine Learning** – Algorithms for Regression (Linear, Logistic), Classification models (Decision Trees, Random forest, XGBoost, SVM, Naïve Bayes, k-NN), PCA, Clustering analysis (k-means, Hierarchical), Forecasting (Moving average, Exponential smoothing, ARIMA, Autoregressive models), Association rules and Recommendation system.

- **Natural Language Processing –** Text Processing, Web Scrapping, Sentiment Analysis, NLTK
- **Deep Learning** – Neural Networks, CNNs.
- **Programming languages**: Python, R and Mysql
- **Excel** and **SPSS**
- **Cloud Platform:** AWS, GCP, AZURE, Heroku
- **Tableau** – Data Visualization, Forecasts, Tables, Charts, Dashboards
- **Big Data** – Basic knowledge of Hadoop, Map Reduce and Spark, Hive

EDUCATION

Education	College	University	Year of pass	% or CGPA
M.Sc. (Agri.) in Agricultural Statistics	College of Agricultural Sciences, Dharwad	University of Agricultural Sciences, Dharwad	2019	9.37
B.Sc.(Ag.MaCo)	Gandhi Krishi Vignana Kendra	University of Agricultural Sciences, Bengaluru	2017	8.11
PUC	Sri Sharadhamba Independent PU Science College	Karnataka state PU board	2013	71.83%
SSLC	Sri Mallikarjuna Swamy Higher Primary School	Karnataka state SSLC board	2011	89.28%

CERTIFICATIONS

- Received certification for successful completion of course in Data Science from Excelr Solutions.
- Received certification for successful completion of Internship Project through emerging technologies firm Innodatatics Inc.

DECLARATION

I hereby declare that the above mentioned information is correct up to my knowledge and I bear the responsibility for the correctness of the above mentioned particulars.

(Yashaswini P)