Sourav Saha

Personal Information

Bangalore, India
souravsaha1605@gmail.com

*8*981782777 - Portfolio - LinkedIn
Live Resume - Kaggle

Career Summary

4 years of professional experience working in Natural Language processing, Data Science and Data Structure. Proposed multiple business solutions to stakeholders as a part of job role and helped developing solution for them.

Professional Experience

Tata Consultancy Services Ltd. - System Engineer   **January 2017 - Present**

- Proposed multiple business solutions as a part of job role and helped developing solution for them.
- Implemented Named Entity Recognition using LSTM algorithm with Tensorflow
- Recommendation problems using NetworkX and experience working in Grakn graph database.
- Tag Extraction using LDA, POS, and from front-end using JavaScript
- Datasets preparation and implementation of classification and document similarity problems
- Relevant search implementation for assets using PostgreSQL.
- Expert ranking for associates.

FusionCharts Technologies Pvt. Ltd., - JS Developer   **July 2016 - December 2016**

- Adding new functionalities to different type of charts and creating new charts based on company requirements. Implementations were made using maths, data structure with JavaScript.

Shoes on loose, - Python Developer Intern   **October 2015 - November 2015**

Hobby Projects
- **Machine Learning Hackathons**
  - Predict Hospital stay duration for COVID patients using CatBoost, GridSearchCV - Rank : 142 / 19K
  - Predict Future Sales using XGBRegressor - Rank : 277 / 8K
  - Twitter Sentiment Analysis using LSTM - Rank : 416 / 15K
  - BigMart Sales Prediction using StackingCVRegressor - Rank : 1726 / 32K
  - Time Series Forecasting of JetRail passengers using Prophet, Holt's Winter, Sarimax - Rank : 399 / 19K

- **Exam Preparation Online (epon.in)** - An online portal to simplify the email digest for their daily learning activities on machine learning, sports and technology. It also features interview experiences and success stories in simplified format.
  - Filter out useful contents, perform tag extraction and broadcast contents relevant to users through email, website and chrome extension.
  - Extract text using tesseract from Images of question papers and segregating them to individual questions after pre-processing.
  - Extract questions from interview experiences, identify unique questions using XGBoost ML algorithm and classify them to corresponding subjects using Ensemble Classification algorithm. Dataset preparation required for performing Ensemble Classification method.
  - Interview experiences to display only required information based on users need like suggestions, technical questions or others
  - 200+ Subscribers, 300+ monthly visits, 2K unique visits before interviews, 1.2K Facebook likes

- **Google Search Recommendation** - Chrome extension to track online activity of user. Visual representation of online activity through charts and recommended links in Google search page based on their past activities.

Certifications

Machine Learning

- Natural Language Processing in TensorFlow
- Neural Networks and Deep Learning
- TCS Machine Learning E2 competency
- Basic Sentiment Analysis with TensorFlow
- Compare Stock Returns with Google Sheets
- Data Analysis Using Pyspark
- Exploratory Data Analysis with Seaborn
- Predict Future Product Prices Using Facebook Prophet
- The Data Scientists Toolbox by Johns Hopkins University
- Machine Learning Foundations: A Case Study Approach by University of Washington

Skills

Python, Tensorflow, Scikit-learn, Networkx, Grakn graph database, Seaborn, PySpark, JavaScript, AngularJS

AWARDS &
ACHIEVEMENTS
- Only student to be selected for job in Fusioncharts from on-campus
- 3rd prize in **BugSmash** coding event in INNOVACION, Tech Fest, IEM Kolkata.
- Student of the Year Performer Award by college

EDUCATION   **Institute of Engineering & Management**, B.Tech., (DGPA) : 8.42 / 10