|  |  |
| --- | --- |
| **Links:**  [Linkedin/santhoshkumar](https://www.linkedin.com/in/santhosh-kumar-choori/)  [medium/santhoshkumar](https://medium.com/@skumarr53)  [GitHub/santhoshkumar](https://github.com/Skumarr53/List-of-projects)  [Personal Website](https://skumar-djangoblog.herokuapp.com/) | Santhosh Kumar  Skumarr53@gmail.com | +91 953 847 2543 |
| **MACHINE LEARNING**  **Regression**  Linear / Ridge / Lasso  ElasticNet / Stacked  Random Forest Regr / GBDT  **Classification**  Logistic / SVM / kNN  Random Forest / XGBoost  **Recommendation**  Collaborative Filtering  Matrix Factorization  **Natural Language Processing**  Text classification  Text generation  **DEEP LEARNING**  ComputerVision  Caption generation  Machine translation | **EXPERIENCE**  **AIR WORLDWIDE** | Flood Research Engineer  Aug, 2016 – Oct, 2019 | Hyderabad, IN |
| [**DATA SCIENCE PROJECTS**](https://github.com/Skumarr53/List-of-projects)  [**Used Car Price Prediction**](https://github.com/Skumarr53/Used-Car-Price-Prediction)   * Built an end-to-end feature transformation and model selection pipeline for predicting the price of used Cars, helping buyers to make an informed purchase. * Gradient Boosting Regressor turns out to be the best model with 0.033 Mean Squared Logarithmic Error (MSLE). * Designed an interactive Web application for model deployment using the **Flask** framework and Hosted on [**AWS**](http://usedcarpricepredict-env.eba-jdefnbzx.us-east-1.elasticbeanstalk.com/) using Elastic Beanstalk service via **Docker** image.   [**Quora Question Similarity**](https://github.com/Skumarr53/Quora-Question-Similarity-Kaggle)   * Built a Classifier that identifies duplicate questions on Quora to enhance the user experience by instantly providing answers to questions that have already been answered. * Created an NLP transformation pipeline for extracting basic, fuzzy, TFIDF, and Word2Vec features, and then trained various ML models. * XGBoost turns out to be the best model with 84% accuracy.   [**Image Caption Generation**](https://github.com/Skumarr53/Image-Caption-Generation-using-Fastai)   * Implemented image caption generation method proposed in **Show, Attend and Tell** paper using **Fastai** framework to describe the content of images. * Pre-trained **ResNet101** model is used for the Encoder and **LSTM** for Decoder. Achieved 24 BLEU score for Beam search size of 5. * Designed a Web application for model deployment using the **Flask** framework.   **Machine Translation - Transformer based**   * Implemented Transformer based machine translation method proposed in **Attention Is All You Need** paper using **Fastai** framework to translate queries French-to-English. * Achieved 58 BLEU score using **BERT** architecture. * Designed web application for model deployment using the **Streamlit** framework. |
| **TECHNICAL SKILLS**  CODING  • Python • SQL • HTML  • JQuery    WEB SCARPING  • Selenium • BeautifulSoup  ML FRAMEWORKS  • Scikit-Learn • Pytorch • Fastai  WEB FRAMEWORKS  • Flask • Django  CLOUD TECHNOLOGIES  • AWS • Heroku  MISCELLANEOUS  • Jupyter Notebook • VS code  • Docker • Linux • MS Office |
| **HONORS**  • 18k+ visitors on my Medium blogs related to Data Science.  • Performance Recognition Award, 2017  • All India 5th Rank in GATE 2016 | **EDUCATION**  **Indian Institute of Science, Bengaluru**  M.E in Water Resources Engineering (Civil) | 2014 – 2016  **University of Agricultural Sciences, Raichur**  B.E in Agricultural Engineering | 2010 - 2014 |
| **INTERESTS**   * Reading books • Blogging • Solo travelling • Algo Trading |