

Kartik Sirwani

SENIOR DATA SCIENTIST

102, Amethyst Baner, Behind Ganraj Mangal Karyalay, Baner, Pune, Maharashtra, India, Pune - 411045

☎ (+91) 7877190455 | ✉ kartiksirwani@gmail.com | 🏠 about.me/kartiksirwani | 📷 kartikwar | 🌐 kartik-sirwani-512716a8/

Summary

Current Senior Data Scientist in start-up company Foghorn. Experienced in building complex deep learning and machine learning models to solve complex problems and use cases.

Skills

LANGUAGES AND TECHNOLOGIES

- Languages : Python, Nodejs, C#
- Frameworks : .Net, Sails, Express
- OS : Windows, Linux, Mac
- Cloud : ECS, Lambda, CloudWatch, IAM, Autoscaling, Security Groups, VPC, Route 53, EC2, SQS, S3, Load Balancer
- DataScience : Python, TensorFlow, Keras, Pandas, Scikit Learn, matplotlib
- Databases : MongoDB, SQL

Work Experience

Razorthink

Bengaluru

SENIOR ARTIFICIAL INTELLIGENCE ENGINEER

Sept 2018 - PRESENT

- Question Answering : This service is used to answer question asked by the end user, using the data provided by the user as a pdf or html document. Did topic modeling using latent dirichlet allocation, this generated topics with word probabilities. Applied a word2vec model on top of it to generate topic vectors for each topic in pdf. Ranking of topics was done by using these topic vectors, these topics were then fed to Bert model for getting answer. Technologies Used: python, nltk, gensim, tqdm
- Table Detection : This service is used to detect tables (bordered and borderless) or table like structures in a pdf document. Built and trained a deep learning model for detecting tables. This was built using Faster Rcnnet(research paper) and curriculum learning. Pretrained VGG16 net was fine tuned on our documents to achieve a precision of 84 percent. Technologies Used: python, tensorflow, opencv
- Template Detection : This service assigns templates to various documents, two documents belong to the same template if they have similar type of text or structure. Built and trained a siamese network with pretrained VGG16 net to detect whether two images are similar or not. If the first page of two different documents were detected to be similar by the siamese network, they were mapped to the same template in the database. Technologies Used: python, tensorflow, opencv, MongoDB
- Carigans : Built a carigan service that converts images of real people to caricatures on the AI Platform of our company. This service was based on a research paper by Microsoft. It makes use of Generative Adversarial networks along with facial recognition to generate cartoon like faces. Gave a presentation on the same (<https://www.youtube.com/watch?v=JeWbg83HZes>). Technologies Used: python, tensorflow, opencv, scikit-learn

Nowfloats

Hyderabad

DATA SCIENTIST

June 2016 - Sept 2018

- Developed Wildfire Acquisition Service. It finds potential customers for buying our product who are then targeted to generate revenue. Achieved a high conversion rate of 20 per cent. Used Machine Learning(Logistic Regression) and Pearson Correlation to train the classifier on previous data. Technologies Used: Python, Scikit Learn, Pandas, Matplotlib, MongoDB, MySQL
- Developed Update Categorization Service. It fetches all the product updates made by the customers on their websites and then categorizes them into offers, discounts, price. Applied a bag of keywords model to classify the updates into categories. Technologies Used: Python, Scikit Learn, Pandas, Matplotlib, MongoDB
- One of the lead developers of NFX. This api project syncs merchant updates, reviews across social channels like facebook, linkedin, twitter, quikr and also creates their online presence on other listing channels generating ROI. NFX has total 19,000 customers and handles around 50,000 updates on a weekly basis. Made REST APIs, services, lambda functions, cron jobs and deployment on ECS. Technologies Used: Python, NodeJs, ECS, Docker, Lambda, Express, MongoDB, Route 53, Ubuntu, SQS, Python

Certificates

Nov 2017 **Applied Machine Learning in Python**, University of Michigan on Coursera

Oct 2017 **Applied Plotting, Charting and Data Representation in Python**, University of Michigan on Coursera

Sept 2017 **Introduction to Data Science in Python**, University of Michigan on Coursera

Education

Birla Institute of Technology and Science, Pilani

Pilani, India

B.PHARM. PRE-PHARMACY STUDIES

2012 - 2016

CGPA - 7.316/10

