Gaurav Nandode

Developing Data-driven Algorithms

India | +91 8172961868 | gsnandode@gmail.com
linkedin.com/in/gsnandode/

Work Experience

Lead Data Scientist — July 2022 - Present

Coditas

- Spearheaded data science initiatives and contribute individually to data science projects while mentoring junior data scientists.
- Developed a mathematical optimization model for bunker planning, reducing costs by 10%.
- Implemented econometric models for behavioral analysis, resulting in 15% fuel savings.
- Worked on SQL based logic for getting accurate Remainder-on-Board number on a ship at any port
- Worked on a sales forecasting problem to estimate expected bookings in advance for a global leader in container shipping to plan the outbound sales using tree-based regression model
- Worked on estimating the gross margin to track the company profits daily for one of the top children's clothing and lifestyle e-commerce companies.
- Estimated the possible returns in real time along with predictions for key business metrics for e-commerce viz. shipping costs and revenue share.
- Build looker dashboards that consume ML model predictions and contributed to refining data engineering pipelines
- Using Recency-Frequency-Monetary scores worked on creating customer segments and build various dashboards for customer journey

Senior Data Scientist — June 2021 - July 2022

AppZen

- Automated invoice processing using Random Forest Classifiers, improving accuracy by 20%.
- Enhanced multilingual support through Neural Network models.

Data Scientist — November 2018 - June 2021

Talentica Software

- Classified logs using MinHash and Locality Sensitive Hashing, improving crash issue detection by 30%.
- Extract key indicators from a crash log viz. key log lines, key function calls, register values, action keywords and crash keywords using Regex.
- For the unclassified logs, applied clustering to create different buckets or unsupervised clusters of logs.
- Also tried extracting Word2Vec, TFIDF, Jaccard Similarity features and applied K-mean and KD-tree clustering methods.
- Implemented a custom K-means from scratch using Jaccard scores.
- Optimized data processing pipelines with PySpark, reducing processing time by 25%.
- Being a part of a product development team, wrote test cases and handled the end to end flow and architecture of the project.

Data Specialist: Cognitive Computing — August 2016 - November 2018

IBM India

1. Cognitive Signature Capture and Verification (03/2018 – To 11/2018)

- Implemented thresholding and segmentation techniques to properly form segments over signatures.
- Worked on extraction of text and signatures within table format by detection of lines using Hough Lines.
- Extracted crucial features from an ensemble of HOG (Histogram of Gradients) and LBP (Local Binary Patterns) features for signature and text segments using PCA (Principle Component Analysis)
- Trained a SVM (Support Vector Machine) classifier on crucial features

About Me

Data Lead with 8 years of experience in developing data-driven algorithms, mentoring junior data scientists, and leading data science initiatives. Proven track record in optimizing logistics, econometric modeling, and sales forecasting for global companies.

Education

- M Tech (Electrical Engineering)
  IIT Kanpur, 2014 - 2016
- B Tech (Electronics & Communication Engineering)
  VNIT Nagpur, 2009 - 2013

Skills

- Python
- SQL
- Data Engineering
- Image Processing and Computer Vision
- Natural Language Processing
- Applied Machine Learning
- Deep Learning

Master's Thesis

Unbounded PML-based Forward Modelling of a 2D GPR Tomography Problem

- Optimization and Implementation of a finite-element-method (FEM) based formulation for a Helmholtz-equation

- Worked on NER (Named Entity Recognition) and Class Recognition for the authorised signatories

2. C3 - Cognitive Content Collator (10/2016 - 02/2018)

- Played a crucial role in implementing an algorithm for image segmentation.
- Trained a SVM classifier to classify image and text segments using HOG features.
- Tried and tested Google and IBM Speech-to-Text conversion tools for the audio chunking part.
- Worked on silence detection to detect pauses in the video and speaker label identification.
- Extracted POS (Parts of Speech) tags from text and script chunks in videos.

3. Auto Insurance Damage Assessment (10/2016 - 11/2016)

- Damage extent classifier based on the confidence score of 3 classes low, moderateand heavy damage: http://visual-recognition-nodejs-gnandode-1616.mybluemix.net/
- Damaged Part classifier to tell which part of the car is damaged: http://visual-recognition-democlassifier2.mybluemix.net/
- Deployed demo in NodeJS environment on IBM Bluemix cloud

modeled forward problem of GPR (Ground Penetrating Radar) tomography using an unbounded function as a perfectly matched layer

Patent

Log analysis in vector space

Patent No. : *US11138059B2*

https://patents.google.com/patent/US11138059B2/en?oq=US11138059B2

Talks

Container Shipment Demand Forecasting Using Regression Models

Global AI Bootcamp