ANURAG BHATT

DATA SCIENTIST, TSCNET Services GmbH, Munich, Germany
PHONE : +49-173 5626609 • E-MAIL : anuraag.bhatt@gmail.com • Homepage : https://anurgbht.github.io/home

To be out there, in the midst of it, endowed with a capacity for leadership, tenacity and enterprise, and a sense of responsibility, delivering the ends and enjoying what I do.

About Me

5+ years of experience in the field of data analytics. Starting from a theoretical statistical thesis in my Master Program to development, execution, validation and deployment of ML models in auto, power, manufacturing, finance and retail industry.

Education

Integrated B. Tech - M. Tech in Mechanical Engineering and Statistical Methods – IIT Kanpur – 2016
CGPA: 8.2 / 10

Master of Arts in Philosophy
IGNOU – 2018 – 77.2%

Skills

Machine Learning, Modelling, NLP

TensorFlow, H2O.ai, PyTorch

Python, R, MatLab, SQL

C#, C++, Java, HTML/CSS/JS/D3

Git, Presentation, Documentation

Academic Publications

- Submitted to archives: A Two-Step Item Response Theory Model for India's Graduate Aptitude Test in Engineering

Academic Conferences

- ICASP-5 & CSSCR-5: The use of Advanced Analytics on Engineered Feature to detect Sticker Breakout in Continuous Casters

Area of Work and Responsibilities

- Leveraging advanced machine learning techniques – ANN, SVM, GBM, Random Forests, Deep Learning – to develop robust predictive models.
- Development and automation of reporting framework pipeline including data storage, analyses, visualization and creating web applications.
- Involvement in direct client engagement, translating complex business problems into data driven solutions.

Work Experience

TSCNET Service GmbH, Munich, Germany
Data Scientist *Oct '19 – Present*

TSCNET Services turns data into information for a secure, prosperous and sustainable Pan-European synchronized power transmission grid.

***Early Warning System*:**
- Predicted overloaded transmission grid elements for mitigative action.
- Extracted relevant features from the available historical data to train nuanced models to predict the overloaded elements.
- Insights generated during the project were leveraged to further enhance data quality and operational understanding.

***Automated Data Quality Assurance*:**
- Automated report generation and consistency checks on all the data streams and pipelines to ensure robust and reliable data.
- Continuous monitoring and outlier detection in the data streams, dashboards and web apps for visualization, and email notifications.

Tata Insights & Quants (Tata iQ), Bangalore, India
Data Scientist *Aug '16 – Aug '19*

Tata Insights and Quants offers multi-sectoral advanced analytics and data engineering solutions using sophisticated predictive analytics and machine learning algorithms.

Unsupervised anomaly detection coupled with supervised classification:
- Used reconstruction error based unsupervised anomaly detectors for accurate tagging of the time series sensor data.
- Unsupervised detection used to generate upstream features that were fed to a supervised classifier.

***ML using sensor data*:**
- Used advanced machine learning techniques for pattern recognition of time series signatures involving innovative feature engineering techniques.
- The solution is currently deployed in Tata Steel, boosting sticker detection accuracy and increasing production.

Genetic Algorithm for vessel lineup optimization:
- We developed a Genetic Algorithm based approach to arrive at a feasible vessel lineup that minimizes the demurrage costs.

Academic/Personal Projects

- Personal tool for language learning using language translation APIs and NLP, NLTK, in Python.
- Custom decision tree tool that supports editing the nodes and splits, in Python.
- Python tool to convert python model objects into machine implementable C++/C# codes. The class can handle GBM, RandomForest, Keras MLP, XgBoost and SVM models. This becomes essential when deploying the code in production environments.
- AutoIt based python script to automate the extraction of the machine generated data
- Genetic Algorithm: Combinatorial Optimization of thickness of a composite plate using GA
- GUI based platform for Digital Image Correlation of plates subjected to strain in MatLab
- Imputation of sparsely filled matrices

Extra-Curricular

- Maintaining a literary blog featuring poetry, short stories and essays on various topics.
- Worked as an executive member with the NGO Svagatagami in the project Abhilasha aiming to help the differently abled kids using technological innovation
- Green - one belt in Tae Kwon-do. Gold medal in district Taekwondo demonstration and Sparring

***Social Media based Lead generation*:**

- Used NLP techniques to mine the content on social media sites like Facebook, Twitter, etc. to identify posts (and users) who may be interested in credit products (like cards, loans, etc.).

Master Thesis, IIT-Kanpur, India
Graduate Student *Mar '15 – July '16*

***A Two-Step Item Response Theory*:**

- Used Item Response Theory in conjunction with Maximum Likelihood Estimation to determine the unique ranks of aspirants and difficulty of questions.
- The results were correlated with the existing GATE score to ensure consistency in the predictions

***Parameter estimation of 3-parameter Fréchet Distribution*:**

- Developed a framework to find all parameters of a 3-parameter Fréchet distribution useful in describing the statistics from Non-Destructive Testing (NDT) of concrete using Ultrasonic Pulse Velocity (UPV) and find the confidence intervals for estimated parameter

GE India Innovation Centre (Aviation), Bangalore, India
Under-graduate intern *May '14 – July '14*

GE India Innovation Centre caters to the design, simulation and analytic needs of the General Electric. GE aviation has designed, developed and patented various turbofan and turboprop engines and related technologies.

***Prediction of Engine Failure*:**

- Used Cox's Proportional Hazard model on the available engine generated data to predict engine failure and flag the problematic engines sufficiently in advance to allow for servicing.
- Engineered features on vibration, pressure, and temperature data through taxi-takeoff-cruising-landing-braking to diagnose preventable ailments in the engine.

Data downloading tool:

- Developed a Java based software to facilitate the downloading of flight data through Oracle/Teradata based servers using a structured graphical interface
- Automated the downloading process using a system scheduler to maintain an up-to-date local database.