



Saluru Durga Sandeep | ME16B125 | PR No : 22/ME/21/125

Indian Institute of Technology Madras

EDUCATION		
Program	URL	Grade
B.Tech in Mechanical Engineering Dual Degree in Data Science	<a href="#">IIT-MGrade Sheet</a>	9.44
XI + XII	<a href="#">Inter Grade Sheet</a>	98.7 %
X	<a href="#">10 Grade Sheet</a>	9.8 / 10
SCHOLASTIC ACHIEVEMENTS		
1. One among the top 30 students selected for <b>Inter Disciplinary Dual Degree in Data Science</b> on a meritorious basis. - <a href="#">Proof</a> 2. Selected for <b>Gartner LEADIT Rotational Program</b> 2020 as an Associate Employee,Day 1 in IITM Internships. - <a href="#">Proof</a> 3. <b>Rank 4</b> in <b>Grocery Sales Forecasting Hackathon</b> conducted by Machine Hack and Overall Global Ranking 62*. - <a href="#">Proof 1,2</a> 4. <b>Rank 37</b> among 6514 participants in Analytics Vidhya JantaHack <b>Mobility Analytics Hackathon</b> . – <a href="#">Proof</a> 5. <b>Rank 40</b> among 5573 participants in Analytics Vidhya JantaHack <b>HR Analytics Hackathon</b> . – <a href="#">Proof</a> 6. Secured a <b>Rank 1636</b> in IIT-JEE Advanced Examination and Rank 661 in JEE Mains Examination 2016. – <a href="#">Proof</a>		
COURSEWORK		
Fundamentals of Deep Learning	Natural Language Processing	Mathematical Foundations of Data Science
Introduction to Data Analytics	Data Analytics Laboratory	Introduction to Data Structures and Algorithms
Feature Engineering for Machine Learning - <a href="#">Proof</a>	Improving Deep Neural Networks – Tuning and Error Analysis - <a href="#">Proof</a>	Introduction to Data Science in Python - <a href="#">Proof</a>
Sequence Models - <a href="#">Proof</a>	SQL Packt Workshop - <a href="#">Proof</a>	Structuring Machine Learning Project - <a href="#">Proof</a>
PROFESSIONAL EXPERIENCE		
<b>Junior Data Scientist Internship @ GYAN DATA [May 2019 – July 2019] – <a href="#">Proof</a></b>		
<b>Furnace Loading Problem @ CUMI</b>	Guided by Prof. Raghunathan Rengaswamy (Dept. of Chemical Engineering, IIT Madras) <b>Packages:</b> CasADi, PuLP, Scipy, NetworkX Solvers: Couenne, Bonmin, CBC, GLPK 1. The Project Statement involved <b>Efficient Packing of Hollow Cylinders</b> in a 3-Dimensional furnace. Basically, a <b>Non-Linear Programming Problem(NLP)</b> to be optimized. 2. The project simplified by breaking into three different problems <b>namely Nesting, Layer Loading and Basket Loading</b> . <b>Mix of Solvers and Graph theory</b> used to get the Optimum. 3. <b>Proposed</b> the idea of <b>Multi-Level Nesting</b> and implemented it using graph theory in their manufacturing process. Was able to increase the <b>furnace efficiency by around 30%</b> .	
<b>ETA and Fuel Modelling in Sea Navigation @ Navidium</b>	Guided by Prof. Shankar Narsimhan (Dept. of Chemical Engineering, IIT Madras) <b>Packages:</b> Ipyleaflets (Geomaps) for Visualisation, StatsModel for Modelling and Analysis 1. The problem involved building <b>ETA and Fuel Model</b> using forecasted <b>weather data</b> . 2. Built an <b>Energy Consumption Model</b> based on <b>Regularised Polynomial Regression</b> with 10 features, Ideal Velocity, wind Speed & direction, course over ground, weight, ship draft, RPM. 3. Built an <b>ETA model</b> (vector) based on ensemble of first principles + Time Series Model. 4. Achieved a Max.& Min. error of <b>0.7% &amp; 0.2 %</b> for the fuel model & <b>5%</b> error in ETA model.	
<b>GRM Program IBM Research Internship @ IBM [May 2020 – July 2020] - <a href="#">Proof</a></b>		
<b>Natural Language Analytics “NLAx”</b>	1. Proposed Problem Statement is that given a <b>Natural Language Query (NLQ)</b> , interpret the query in analytics scope and produce insightful results. Existing QA can only solve point queries. 2. I am working on creating an independent <b>interface between NLQ and SQL</b> which is known <b>Ontology Query Language a.k.a OQL</b> . Building the language using <b>Antrl 4.0</b> . 3. By this Project, we can solve <b>90% of the Online Analytical Processing (OLAP) queries/KPI</b> meant for business users.	



Saluru Durga Sandeep | ME16B125 | PR No : 22/ME/21/125

Indian Institute of Technology Madras

Data @ ANZ Virtual Internship Program – InsideSherpa – <a href="#">Proof</a>	
<b>Predicting Salary From Transactional Data</b>	<ol style="list-style-type: none"> <li>1. Rigorous Exploratory Data Analysis (<b>EDA</b>) and found some key-insights like when the transactions are heavy and less based on seasonality factors, Identifying the potential customers using <b>RFM model</b> which can be used for Marketing Strategies.</li> <li>2. Ensemble model to predict the salary of the customers. Results were not reliable because of less availability of data.</li> </ol>
Dual Degree Project – Quantel.Ai - Proof – This will be updated in around September	
<b>Management Discussion and Analysis (MD&amp;A) Analytics</b>	<ol style="list-style-type: none"> <li>1. MD&amp;A, this gives the information of company's forward lookings and expected behaviour for future. This is used in investing in the potential companies.</li> <li>2. Built a <b>language model Financial BERT</b> which is trained on 2002 to 2016 10K- SEC filings.</li> <li>3. <b>SEC Filings report summarisation</b> based on research standard keywords of 17 topics and corresponding <b>sentiment of the sentences using FinBERT</b></li> <li>4. Further project work involves fine grained analysis of whole SEC report and <b>Portfolio Optimisation</b></li> </ol>
Course Projects & Hackathons	
<b>Grocery Sales Prediction @ Machine Hack</b> <a href="#">Code Link</a> <a href="#">Leaderboard Link</a>	<ol style="list-style-type: none"> <li>1. One of the Top 5 contestants in this hackathon (Rank 4). This hackathon is about predicting future sales based on the past sales but <b>without any info on date</b>.</li> <li>2. Converted a <b>Time series data to Time Invariant data (stationary data)</b> and built a robust model which is stacked ensemble model of LGBM, XGB, Bayesian, Random forest with Extra trees as Meta model. Simple but stood-out in Top of LB.</li> </ol>
<b>JantaHack Mobility Analytics @ Analytics Vidhya</b> <a href="#">Code Link</a> <a href="#">LeaderBoard Link</a>	<ol style="list-style-type: none"> <li>1. One of the Top 50 contestants in this hackathon (Rank 37). This hackathon is about prediction surging price type based on customer behaviour.</li> <li>2. Quite challenging because of lot of missing values in the dataset and the given features are not informative. So did manually <b>feature engineering like target encoding and interactive feature encoding</b>, built a strong cross validated LGB model.</li> </ol>
<b>Deep Learning Project @ CS6910</b> <a href="#">Grade sheet Link</a>	<ol style="list-style-type: none"> <li>1. Built a <b>Stacked Auto-encoder from scratch in torch</b> for <b>Dimensionality Reduction</b> and <b>noise removal</b>, gave a reduction in size of <b>more than 80%</b>.</li> <li>2. Built an ensemble of pre-trained models from torch for classifying the type of dress using <b>MNIST Fashion Data (Fine-tuning also included in ensemble)</b></li> <li>3. Built a <b>Seq2Seq Neural Machine translation</b> model which <b>translates English to Telugu</b> with <b>accuracy of more than 95%</b>. Constrained by data to improve.</li> </ol>
<b>Zimnat Insurance Recommendation Challenge @ Zindi</b> <b>Participation <a href="#">Proof</a></b>	<ol style="list-style-type: none"> <li>1. Predicted which insurance products existing clients will purchase next based on other clients purchases with a <b>log-loss of 0.037, AUC-score of 99.1% and balanced accuracy of 98.5%</b></li> <li>2. Handled <b>Imbalanced dataset</b> with 95% without purchases and 5 % with purchases using data balance techniques like <b>Over-Sampling using SMOTE, Under-Sampling using Tomek Links</b>.</li> <li>3. Converted recommendation problem into classification problem and applied an ensemble of three top boosting models (XGB, LGBM, Catboost) with k-fold cross validation</li> </ol>
<b>Air Quality Forecast Challenge @ Zindi</b> <b><a href="#">Participation Proof</a></b>	<ol style="list-style-type: none"> <li>1. Predict the air quality level at exactly 24 hours after a 5-day series of hourly weather readings from 5 different sensors with <b>RMSE score of 34.98 PM (Particulate Matter)</b>.</li> <li>2. Extensive feature engineering to get <b>200+ features</b> like aggregation features, target encoding without data leakage, lag features, variable transformations.</li> <li>3. Model is a stacked ensemble using LGBM as base model and CatBoost model as meta-model</li> </ol>
Position of Responsibility (POR)	
<b>Deputy Placement Co-ordinator - Placement Team IIT Madras</b>	<ol style="list-style-type: none"> <li>1. Organised campus placement drive for all the pre-final year students in Mechanical Engineering department of IIT Madras.</li> <li>2. Confirmed over 30 companies for the campus placement season 2018-2019.</li> <li>3. Coordinated and planned the campus placement process of over 20 students from the department of Mechanical Engineering.</li> <li>4. Coordinated with over 10 companies during the campus placement interviews.</li> </ol>