SIDHANT THOLE

◇ Github:SPThole ◇ Google Scholar ◇ tholesidhantp@gmail.com ◇ +91-9762112543
Linkedin

EDUCATION

Indian Institute of Technology Madras *Chennai, Tamil Nadu, India*
MS by research in Engineering Design (8.35/10.0) *July 2017 - July 2020*
Thesis: **Interpretable Self Organizing Maps for Design Space Exploration and Optimization**
Courses: Machine Learning for Engineering and Sciences, Applied Statistics, Engineering Optimization

Vishwakarma Institute of Technology Pune, Maharashtra, India
B.Tech. in Mechanical Engineering (8.81/10.0) *July 2013 - July 2017*

PROFESSIONAL EXPERIENCE

Data Science Associate, ZS Associates, Pune, India *August 2020 - Present*

- **Competitive Insights Engine: Modelling perceptions to actions**
  - Created efficient data processing pipeline in order to ingest variety of survey and sales data coming from multiple sources
  - Built an **automated modelling tool** to predict sales from survey data using hyperparameter optimization module
  - Implemented insights tool which gives out strategic insights using **Shapley Values** in **50%** reduced time

- **Predictive Insights: Customer centric marketing**
  - Developed a deep learning model to predict if the customer will increase or decrease the sales with f1 score of ~ 60
  - Generated **customer level insights** that would be helpful for sales representative in sales call

- **Auto-binning: Optimized bins for univariate analysis**
  - Developed a novel solution for auto binning the continuous features using **multi-objective optimization**
  - Devised solution carry out binning of the continuous feature to capture interesting trends
  - Tool **reduced time spent** on EDA, missing value imputation, variable encoding by **30%**

- **Xplainable Models: Module to explain predictions by machine learning models**
  - Built an **explainable-AI module** that takes any ML model and explains the predictions using various methodologies
  - Module gives out **feature attributions** along with confidence intervals for them

- **Brand Launch Analytics: Boosting pharma sales and marketing using machine learning**
  - Identified **key drivers** to achieve launch success of 4 drugs for top pharma clients in different therapeutic areas
  - Generated customer target list using adoption propensity modelling, which reduced efforts of sales team by $\sim 20 - 30\%$
  - Gave **actionable insights** to clients using what if scenarios based on Model Based Recursive Partitioning

PUBLICATIONS AND PROJECTS

- **[Journal Paper] Design space exploration and optimization using self-organizing maps** *July 2020*
  - Developed novel neural network algorithm for **multivariate response surface visualization**
  - Identified **regions of interest** for optimization and thereby achieved optima with $\sim 40\%$ **fewer data points**
  - Designed algorithm used for visualizing patterns in data in a **single plot** and make **informative decisions** on modelling (Sampling, EDA, correlation hunting, outlier detection) and optimization

- **[Conference Paper]Identifying robust design spaces using Self Organizing Map** *WCSMO 13- Beijing*
  - Recognized regions of optima which are robust to uncertainties in input features *May 2019*

- **Kaggle BMS – Molecular Translation (ZS Associates Team)** *May 2021*
  - Predicted InChI sequences for a molecule present in scanned **noisy** image ($\sim$ **2.4M** labeled images)
  - Achieved mean Levenshtein distance of **1.43** over **1.6 M** test images, ranked 56th out of total 874 teams participated
  - Implemented image captioning solution using vision models and LSTM with attention mechanism

- **Github Other projects**
  - English to Hindi transliteration using LSTM with attention
  - Sentiment analysis using BERT model trained on google playstore comments

TECHNICAL STRENGTHS

Python   scikit-learn, pandas, numpy, scipy, PyTorch, matplotlib, seaborn, boosting algorithms, optuna
Others   MATLAB, AWS (EC2, S3), LaTeX, Google Colab, Jupyter, MySQL

POSITIONS OF RESPONSIBILITY

IIT Madras Placement Team : Coordinated organization of placement activities for $\sim$**1300** students and $\sim$**300** companies
Teaching Assistant : Served as TA for Optimization methods in Engineering Design course at IIT Madras and other NPTEL courses