# Kuldeep Singh Dahiya

Final Year Undergraduate
Department of Mechanical Engineering

: ksdahiya@iitk.ac.in
: +91 7419196581
: Kuldeep Dahiya : Kuldeep Dahiya

## Academic Qualifications

| Year | Degree/Certificate | Institute | CPI/% |
|---|---|---|---|
| 2019 - Present | B.Tech | Indian Institute of Technology, Kanpur | 6.63/10 |
| 2018 | CBSE(XII) | Nalanda Vidya Niketen Sr. Sec. School,Dalanwas | 92.6% |
| 2016 | CBSE(X) | Nalanda Vidya Niketen Sr. Sec. School,Dalanwas | 10 CGPA |

## Programming Achievements

- Qualified to **Round 1** in Global Competitions **Google Codejam 2022** and **Facebook Hackercup 2021**.
- Secured a **Rank 16** among **267** Registered Teams in **Shaastra Programming Contest Finals** hosted by **IIT Madras**.
- Achieved a **Global Rank of 126** among over 8000 participating candidates in **Codechef** Starters 43 Divison 2.
- Achieved a **Global Rank of 313** among over 20,000 participating candidates in **Codeforces** Round 800.
- Rated **Expert** Programmer on **Codeforces** with a rating of **1649** and **5*** on **Codechef** with a rating of **2050**.

## Research Internship

- **Samsung Research Institute Bangalore** | Research Intern *(May'22 - Jul'22)*
  - Re-iterated over stages with some variations to improve the performance of **CLA(Closed-Loop-Automation)**. Especially inclusion of cellNumber in **stitching** and other variations like changing the boundaries etc.
  - Finding the root cause behind a **network degradation** and improving its accuracy or performance.
  - Loaded KPI's data/hourly aggregated, Derived **Diamond KPI** (Pre-Processing), Domain Labelling, Model Training using **XG-Boost** anomaly detection model and Model Validation.
  - Used **Z-Score normalization** to normalize the Diamond KPIs for all combinations of eNodeB and cellNumber.
  - Stitched **eNodeB + CellNum** as a tool and then doing end to end performance of **CLA**.
  - Got **Pre-Placement Offer(PPO)** from **Samsung Research Institute, Bangalore** for excellent performance in the internship.

## Key Projects

- **Credit Card Fraud Detection** | **Self Project**
  - Built a model to classify fraud transactions dealing with **high-class imbalanced datasets** containing transaction data.
  - Applied **PCA** for dimensionality reduction, **t-SNE** for data visualization, and **InterQuartile Range** for outlier removal.
  - Performed various techniques like **Random Undersampling, NearMiss,** and **SMOTE** to handle class imbalance.
  - Applied various ML algorithms including Logistic Regression, SVM, KNN, and decision tree, achieving a recall of **95.0**%.
  - Compared different models based on their **F1 score** and **Confusion matrix** and plotted their **ROC** and Learning Curves.
  - Increased F1-score to **74.0**% by hyperparameter tuning and varying the final ratio of different classes in **SMOTE**.
- **Football Analysis: xG Model** | **Self Project**
  - Created an **Expected Goals model**: Predict whether or not a shot will be converted into a goal.
  - **Scraped** events data and created attributes in a cylindrical coordinate system for **open goal classification**.
  - Applied **Logistic Regression** to find **probability** based on **goal angle** at a shot location and shot distance.
  - Visualised and presented the calculated **xG value** on the football field for validation and further observatory analysis.
  - **Prospects**: Set-piece specific classification, detailed real-time attributes, players' expertise for further improvement.
- **Sentiment Analysis of Reddit Comments– Machine Learning** | **Self Project**
  - Analyzed data consisting of Political posts of 5000 Data-points with 21 variables like Audience, Message, Political party.
  - Used **TF-IDF** and feature Engineering to preprocess data and convert it into a compatible input dataset.
  - **K-Means clustering** was used over the **TF-IDF matrix** to classify variables into 10 clusters based on document similarity.
  - **KNN** algorithm was used to train the model for classification of comments into topics of interest.
  - Testing set **F1 scores** were recorded at **75** and **71** and achieved R2 score of **0.82** implying the reliable accuracy of model.
- **Restaurant Management System** | **Self Project**
  - Implemented a back-end algorithm for ordering food items in a systematic way for the users and calculating bills.
  - Developed in C Style using **Pointers** to allocate memory dynamically to store food orders and used function for statistics.
  - Used **Linked List Data Structure** to dynamically store and optimize operations needed in ordering and canceling food.

## Technical Skills

- **Programming Languages:** C, C++, SQL, Python
- **Machine Learning:** Tensorflow, Pandas, Numpy, Matplotlib, Seaborn, Keras, sci-kit learn, XGBoost
- **Softwares and Libraries:** MS Office(Excel, Powerpoint, Word), GNU Octave, Micro-Cap, AutoCad, LaTeX

## Extra-Curricular Activities

- Skilled Lawn Tennis as a part of necessary physical exercise.
- Participated in Hall-5 Parade in Galaxy-20 (Inter-hall Extra-curricular event).
- Solved more than **3500** problems on **world-renowned platforms** like Codechef, Codeforces, and Leetcode.