

# KEERTHANA E

E-mail id: [keerthanaeganathan@gmail.com](mailto:keerthanaeganathan@gmail.com)

LinkedIn: <https://www.linkedin.com/in/keerthana-eganathan-39396a170/>

Mobile: +91 8610711564

## Career Objective:

Qualified Machine Learning Engineer adept with big data Storage, Processing and Computation. Highly skilled with ML/ AI – Data Science Domain. Experienced Collaborating with cross functional teams to deliver cutting-edges solutions.

## Experience:

**Machine Learning Engineer**, UGDx – Formerly known as INSOFE - Bangalore. (NOV,2022 - till present)

**Trainee Data Scientist**, INSOFE – Bangalore. (Dec,2021- Oct,2022)

## Educations:

Master Of Engineering in Computer Science 2020 | 8.86 CGPA | Priyadarshini Engineering College | Anna University  
Bachelor Of Engineering in Computer Science 2018 | 7.68 CGPA | Priyadarshini Engineering College | Anna University  
Higher Secondary Certificate HSC 2014 | 72% | St paul's Matric Higher Secondary School  
Secondary School Certificate 2012 | 75% | St paul's Matric Higher Secondary School

## Certifications:

Completed Post Graduate Program in Computational Data Science from Insofe – Bangalore. (Jan,2022 - Oct,2022)  
The Program is certified for quality, pedagogy and assessment by Case Western Reserve University.  
Data Science Certification at NUCOT – Bengaluru.

## Skills:

- Python (Pandas, NumPy, Scipy ),SQL, Basics of R, Advance MS-Excel, Power point.
- Fundamentals of Data Structure
- Visualization tools: Matplotlib, Seaborn, Basics of Power BI.
- Exploratory Data Analysis, Feature Engineering, statical Modeling
- Learning Algorithms: Supervised and unsupervised Algorithms.
- Improving Learning Models: Grid search, Randomized Search, Hyper Parameter Tuning.
- Big Data Engineering: Linux, Hadoop, Pyspark, Hive, Pyspark Machine Learning Pipeline.
- Version Controls: Git and Git-Hub, Basics of Docker and Containerization, Basics of MLOps & DevOps.
- Deep Learning Algorithms – TensorFlow, PyTorch, Keras, CNN, Computer Vision, Natural Language Processing, Hugging Face.

## Projects:

### 1. Developed a classification model to predict the likelihood of hospital readmission for patients diagnosed with diabetes. The objective was to identify high-risk patients who have a higher probability of being readmitted within a specified time frame.

- Utilized a large-scale dataset consisting of 66587 of records and 49 features, including various features such as time\_in\_hospital, medical\_specialty, num\_lab\_procedures, num\_medications, number\_outpatient, number\_emergency and number\_diagnoses. Conducted extensive data preprocessing, including handling missing values, normalizing numeric features, and encoding categorical variables.
- Implemented Decision Tree classifier, KNeighbors Classifier, Random forest in Python, leveraging scikit-learn library to handle the classification task.
- The developed model achieved an overall accuracy of 75%, with a precision of 72% and a recall of 73% in Random forest & Decision Tree. Indicating a reasonable ability to discriminate between patients who would be readmitted and those who would not.
- Additionally, optimize the model's performance using Randomized Search CV and Grid Search CV.
- The prediction model significantly improves patient care by enabling early identification of individuals at a higher risk of readmission. This allows healthcare providers to intervene proactively, implement targeted care plans, and allocate appropriate resources to prevent unnecessary hospital readmissions.

### 2. Developed a classification model to predict the probability of insurance policyholders being fraudulent for an

**insurance provider. The objective was to assess the likelihood of fraudulent activities and identify high-risk policyholders to mitigate potential losses.**

- Utilized a large-scale dataset consisting of 28,836 records with 42 features including the information of Claim data, Demographics data, Policy data, Vehicle data. Conducted extensive data preprocessing, including handling missing values, normalizing numeric features, and encoding categorical variables.
- Implemented KNN classifier in Python, leveraging scikit-learn library to handle the classification task. Additionally, applied cross-validation techniques to ensure robustness and optimize the model's performance.
- The model demonstrated high precision 76.1% in detecting fraudulent one, significantly reducing false positives and minimizing potential losses for the company.

**3. Developed and implemented the Image Classification Model using Convolution Neural Network. The objective is to detect the emotions of the image and classify whether it is sad / Happy.**

- Utilized images from Google (web scraping) 340 real time images in Google classified into 2 classes – Happy & Sad. Conducted extensive data preprocessing, Normalizing the Images on prescribed range.
- Built a Customized Convolutional neural network (CNN) to generate the feature map from those images and flattened the feature map output to feed those output for Customized Artificial Neural Network (ANN) to generate the customized output.
- The developed model achieved an overall accuracy of 65%, with a precision of 60% and a recall of 60% in Convolutional neural network (CNN).
- Testing with real time images generated day to day and seeing the predictions whether it is good enough detect the images emotions.

**4. Developed and implemented the Spam/Ham mail Classification using Naïve Bayes Model. The Objective is to detect the Ham/Spam for the given samples of mails.**

- Utilized the samples of mail for training and testing the Naïve Bayes model. Conducted extensive data preprocessing, including handling missing values, normalizing numeric features, and encoding categorical variables.
- Taking the probability of that token occurs in the given mail is Spam & Non-spam (ham mail).
- Testing the mails and classifying them either it can be ham/spam mail.

**Project Links:**

<https://github.com/keerthanaE702/Impact-on-meditation-for-the-lifestyle-Disease-on-hospital-readmitted.git>

<https://github.com/keerthanaE702/Fraud-claims.git>

<https://github.com/keerthanaE702/Identify-The-Best-Model-For-Class-Imbalance-Data-in-Multiclass-Problem.git>

<https://github.com/keerthanaE702/spam-Ham-mail-classification.git>

**Additional Activities:**

Captivated and engaged students through innovative hands-on sessions on machine learning concepts.

implemented and assessed cutting-edge machine learning models on complex data sets.

Conducted thorough research and data analysis, applying statistical techniques to interpret data and translate findings into actionable insights.

I certify that all information stated in this resume is true and to the best of my knowledge. I authorize any organization to verify the information provided in this resume.

- Keerthana Eganathan.