**Software Requirements Specification**



**Prepared by Sukru Bora Karakus**

**and Ismail Eza for the Project Classifying MR images and Detecting Tumors using ANN’s in Image Processing, and The Contribution of Ensemble Techniques to Accuracy Rate**

Assoc.Dr. Ferdi Sönmez – Advisor

**Team Members**

Şükrü Bora Karakuş (B1905.090021)

İsmail Eza (B1905.090045)

**Istanbul Aydin University – Department of Software Engineering**

**SEN423 Graduation Project – I**

**Fall 2023-2024**

**Table of Contents**

1. **Introduction**
   1. **Problem Definition**
      1. Optimizing Summary Significance in Medical Context
      2. User-Selectable Summary Length for Medical Insights
      3. Configuring a Fast-Responding Server System for Immediate Insights
   2. **System Overview**
      1. Components
   3. **Definitions, Acronyms, and Abbreviations**
   4. **Assumptions and Dependencies**
      1. Assumptions
      2. Dependencies
2. **Overall Description**
   1. **Product Functions**
      1. Use-Case Model Survey
      2. Actor Survey
   2. **Interfaces**
      1. User Interfaces
      2. Software Interfaces
      3. Communications Interfaces
   3. **Constraints**
3. **Specific Requirements**
   1. **Functional Requirements**
      1. Dataset Collection and Preparation
      2. ANN Architectures
      3. Ensemble Models
      4. Model Training
      5. Evaluation Metrics
   2. **Non-functional Requirements**
      1. Performance
      2. Security
      3. Usability
      4. Reliability
      5. Maintainability
   3. **Data Requirements**
   4. **Interface Requirements**
4. **References**
5. **Introduction**

This document is a Software Requirements Specification SRS for the project named “Classifying MR Images and Detecting Tumors using ANN's in Image Processing and The Contribution of Ensemble Techniques to Accuracy Rate”. It outlines requirements that follow common requirements practices as an introduction part, defining goals scope definitions abbreviations acronyms references overview.

* 1. **Problem Definition**

In the field of medical imaging, there is a need for extracting valuable information from large MRI datasets when it comes to healthcare practitioners and researchers in this regard. But the problem emerges as users try to spend less time interpreting these complex images. This issue is tackled by our project through the development of brief MR images summaries facilitating quick information extraction. The key goals of our project tailored for MR image classification include:

* + 1. **Optimizing Summary Significance in Medical Context**

It means that the generated summaries should be highly relevant and cover as many critical aspects of MR images as required by medical diagnosed. The aim is to represent the point of most critical information, bringing in accordance with medical context and purpose behind making these images.

* + 1. **User-Selectable Summary Length for Medical Insights**

Enabling users, in particular healthcare practitioners, to tailor the length of summaries according to their specific diagnostic requirements.

* + 1. **Configuring a Fast-Responding Server System for Immediate Insights**

Establishing a server system that is fast and efficient in providing healthcare personnel instantaneous access to summarised MR image details. Swift decision-making in medical settings depends on minimizing wait times.

* 1. **System Overview**
     1. **Components**
* Image Preprocessing Module

The Image Preprocessing Module is also a module that prepares MR images for analysis.

* Artificial Neural Network (ANN) Architecture

The ANN Architecture of the core part in a system is also responsible for training to learn intricate patterns within MR images. The layers of neurons in the network allow it to extract relevant features that are significant for correct classification, such as tumor detection.

* Tumor Detection Module

This module, therefore, helps the system improve its ability to correctly identify and classify tumors using learned features.

* Model Training and Validation

The training process for the ANN is to present it with a labeled dataset of MR images which will allow it to learn and set its parameters in order correct classification. Validation helps in making the model generalized for unseen data.

* Evaluation Metrics Module

The Evaluation Metrics Module is concerned with the evaluation of a classification model. Accuracy, Precision, Recall and F1 score are metrics that give a holistic picture of the competence in tumor detection by the model.

* 1. **Definitions, Acronyms, and Abbreviations**
* MR : Magnetic Resonance Imaging
* ANN : Artificial Neural Network
* Data Preprocessing : The initial step, in preparing the data involves normalizing, resizing and reducing noise in MR images.
* Tumor Detection Component : A module designed to identify and locate tumors within MR images.
* Model Training : This phase involves the Artificial Neural Network (ANN) learning from datasets to optimize its parameters.
* Evaluation Metrics Component : A module for evaluating the performance of the classification model using metrics such as accuracy, precision, recall and F1 score.
* Classification : The process of assigning a category or label, to MR images based on learned patterns and features.
* Precision : A metric that measures the accuracy of predictions made by the model.
* Recall : A metric that measures the models ability to capture all instances in the dataset.
* F1 Score : The mean of precision and recall providing a balanced assessment of the models performance.
* Accuracy Rate : The accuracy rate represents the proportion of instances that were correctly classified compared to the number of instances, in a dataset. It serves as an indicator of how effective a classification model's, in accurately predicting outcomes.
* Ensemble Techniques : Techniques that aim at improving the accuracy of results in models by combining multiple models instead of using a single model.
  1. **Assumptions and Dependencies**
     1. **Assumptions**:
     + **Quality of the Dataset**

In terms of dataset quality, it is supposed that the MR image dataset used for training and validation is highly rated, properly labelled and representative of variety in medical conditions.

* + - **Training Data Representativeness**

The assumption makes sure that the training dataset is comprehensive in terms of diversity of MR images encountered in real-world medical scenarios. Thus, the model can be generalized.

* + - **Hardware Infrastructure**

The project assumes that there is access to a powerful computing infrastructure with GPUs which are necessary to speed up neural network training.

* + - **Models’ Accuracy Rate**

Project’s overall professionalism and its potential impact to medicine will be enhanced if targeted accuracy rate for the models can be achieved. This may require processes such as applying ensemble techniques so as to increase the accuracy rate.

* + 1. **Dependencies**:
* **Deep Learning Libraries**

Dependence on deep learning libraries like TensorFlow or PyTorch is made in this project for implementing and training the neural network models.

* **Image Processing Tools**

For image preprocessing among other things such as feature extraction, this project depends on image processing libraries and tools like TensorFlow.

* **Data Set Availability**

The project depends on the labelled dataset for training and testing purposes.

* **Model Evaluation Metrics**

Models have dependencies on metric libraries (such as scikit-learn) to evaluate their performance and to understand accuracy rate before applying Ensemble.

* **Python Environment**

The project depends on the Python environment for libraries and frameworks in using machine learning and deep learning applications.

1. **Overall Description**
   1. **Product Functions**

The most important function of the system is to use ANN architecture to detect and classify tumours in the image. It includes components for image feature extraction, tumour detection, model training and evaluation metrics and provides comprehensive functions for MR image analysis. It also improves the current success rate by using the trained models through ensemble techniques.

* + 1. **Use-Case Model Survey**

Use-Case 1 : Classify MR Image

Actor : Healthcare Professional, Researcher

Description : The first use case classifies MR images using a developed and trained Artificial Neural network model. The user uploads the MR image using an interface or through the console and the system first processes the image, extracts its features and passes it through various steps for tumour detection. The system then presents the results showing the type of tumour or absence of tumour if a tumour is detected in the MR image.

**Flow of Events**

1. User accesses the system from the console.

2. User uploads MR image for classification and detection.

3. The system takes the input to the preprocessing step.

4. Its features are handled by ANN.

5. The tumour in the MR image is detected by the ANN architecture.

6. The system performs the classification process.

7. Presents the results to the user.

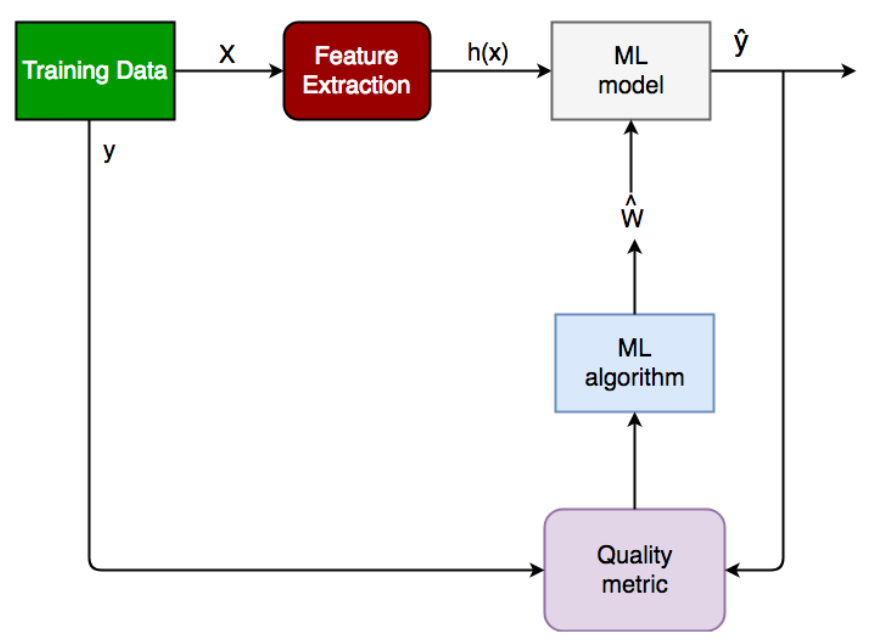


Figure-1 : Classification Diagram usign ML

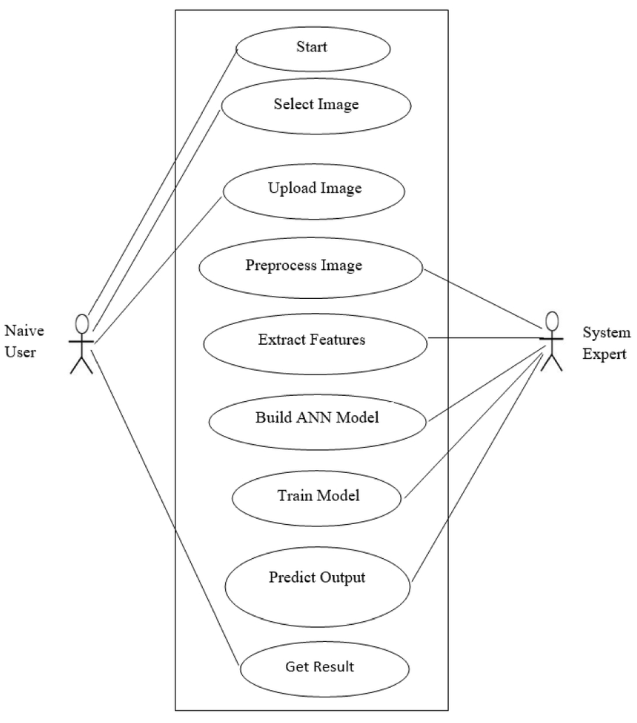


Figure-2 : Classify MR Image Use Case Diagram

* + 1. **Actor Survey**
* Actor 1 : Healthcare Professional (Native User)

Description : Healthcare personnel connect and use the system for diagnosis to ensure accuracy in MRI images and perform tumour detection.

Responsibilities : Upload MR images to the system for classification and analyse the results. Provide feedback to improve the system.

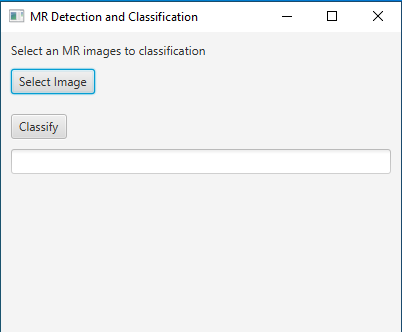
* Actor 2 : System Administrator

Description : Configures and maintains the system. Ensures that the performance is running at optimal values.

Responsibilities : Ensures that the system runs reliable. Improves the system according to feedback from users.

* 1. **Interfaces**
     1. **User Interfaces**

**A simple interface will be created in Java for this application. Users will be able to use it without difficulty because it is a simple interface to use and understand.**



* + 1. **Software Interfaces**
* **Deep Learning Libraries**

In ANN applications, deep learning libraries such as TensorFlow are used for efficient model training and this project is based on the TensorFlow library.

* **Image Processing Tools**

The system works in integration with the TensorFlow image processing library to improve the quality of the inputs, scaling, pre-processing and feature extraction when necessary.

* **Web Technologies and Development Environment**

The application was written on the Jupyter Notebook IDE of the Anaconda application. Jupyter Notebook can run in any local browser and is cross-platform compatible.

* + 1. **Communications Interfaces**
* **User- System Interaction**

Users will connect to the system from a web-based system using .py extension files. This indicates that the system uses web communication protocols (http/HTTPS).

* **External Database**

There is no relational database as a database. Images (approximately 10.000 images data) are stored in a folder in jpeg. It is kept in 4 different classes in jpeg format. It is divided into test and training data.

1. **Specific Requirements**
   1. **Functional Requirements** 
      1. **Dataset Collection and Preparation**

* **Data Retrieval**

The system should allow users to upload MR images for verification purposes.

* **Data Preprocessing**

The system processes MR images for training, optimizing them through preprocessing steps such as normalization and resizing if necessary.

* + 1. **ANN Architectures**

The system develops multiple ANN architectures for MR image classification and applies Ensemble techniques to combine their results. Each ANN architecture contributes to the classification and detection of MR images with its internal layers.

* + 1. **Ensemble Techniques**

The system applies Ensemble techniques on the models to enhance accuracy.

* + 1. **Model Training**

The system transforms itself into a trained model through a learning mechanism on a labeled dataset containing MR images, optimizing ANN parameters.

* + 1. **Evaluation Metrics**

The system provides performance metrics like accuracy, precision, and F1 score to evaluate results from trained models.

* 1. **Non-Functional Requirements**
     1. **Performance**

The system must not exceed the maximum specified time for classification tasks. (The exact time is not determined as models are currently in the training stage and may vary based on user hardware.)

Model accuracy should not fall below the desired/required rate.

* + 1. **Security**

The system should only use and retain MR images uploaded by users for model purposes, refraining from using or sharing any personal information associated with the MR images.

* + 1. **Usability**

The system should provide ease of use for users, even if they lack expertise in software or project requirements.

* + 1. **Reliability**

The system must consistently perform its tasks, maintaining optimal system health with minimal crashes or interruptions.

* + 1. **Maintainability**

The system should be modular to facilitate future updates and maintenance, ensuring openness to development.

* 1. **Data Requirements**

**Dataset Formats**

The system should accept data in jpg, png, and jpeg formats. It should be adaptable to other common MR image formats (such as DICOM) when needed.

* 1. **Interface Requirements**

**User Interaction**

The system should offer comfortable interaction for users to upload MR images and interpret results.

1. **References**

* https://www.linkedin.com/pulse/srs-document-sample-diwakar-singh-/
* [https://www.cse.msu.edu › SRSExample-webapp](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiLx8CfjtGDAxX2SPEDHULiA1gQFnoECA4QAQ&url=https%3A%2F%2Fwww.cse.msu.edu%2F~cse435%2FHandouts%2FSRSExample-webapp.doc&usg=AOvVaw249wPeutmbLpdGGSWXPd77&opi=89978449)
* https://www.reqview.com/doc/iso-iec-ieee-29148-srs-example/
* https://www.analyticsvidhya.com/blog/2020/08/top-4-pre-trained-models-for-image-classification-with-python-code/
* https://roboflow.com/models/classification
* https://jonascleveland.com/best-image-classification-models/
* https://www.researchgate.net/figure/Use-case-diagram-of-the-proposed-model\_fig1\_359271874 (Use case diagram)
* https://bestsucksm.life/product\_details/8747428.html (Classification Diagram)