## Customer Requirements Speciﬁcation

**PO\_SAG\_CRS\_ML\_SR ــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــ**

**Table of Contents**

[Document Status](#_heading=h.gjdgxs) 1

[Document History](#_heading=h.30j0zll) 2

[Reference Document](#_heading=h.1fob9te) 2

[Project Description](#_heading=h.3znysh7) 3

Features 3

Key Elements 3

System Context4

CRS Requirements4

# Document Status:

|  |  |  |
| --- | --- | --- |
| **Name** | PO\_SAG\_CRS\_ML\_SR | |
| Version | V1.0 | |
| Status | proposed | |
| Author | Alaa Eid | |
| Date | [10-11-2022] | |
| Team approval |  |  |
| Mentor approval |  |  |
| Final approval |  |  |

# 

# Document History:

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Author** | **Date** | **Change** |
| 1.0 | Alaa Eid | [10-11-2022] | * Initial Creation * Add definition * Add features * Add Key elements * Add System Context * Add CRS Requirements |

# Reference Document:

|  |  |  |  |
| --- | --- | --- | --- |
| **Ref.number** | **Doc.Name** | **Version** | **Status** |
| 1 | PO\_SAG\_CR\_ML | V1.1 | Released |

# Project Description:

## Definition:

The project aims to help the blind using computer vision to help them identify different things around them and NLP to interact with the user using voice commands.

The project simulates the natural vision of humans using AI techniques enabling the blind to function more normally and independently.

## Features:

* Speech recognition

## Key Elements:

* The model shall transform the audio frames into Mel Spectrograms.
* The model shall encode character label for each audio sample into some numerical value.
* The model shall have 2 architectures of ResCNN and Bidirectional-GRU.
* The ResCNN shall process each audio to extract the features.
* The Bidirectional-GRU shall process the audio features step by step, making a prediction for each frame and output probability matrix for characters.
* The CTC decoder shall take this probability matrix and extract the highest probability characters .
* The model shall return the output to ECU.

## **Speech Recognition**

## System Context:

**ECU**

**Speech\_Recognition**

**Input\_Audio\_SR**



**Output\_Text\_SR**

## 

## CRS Requirements:

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***001***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The model shall transform the audio frames into Mel Spectrograms. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***002***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The model shall encode character label for each audio sample into some numerical value. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***003***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The model shall have 2 architectures of ResCNN and Bidirectional-GRU. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***004***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The ResCNN shall process each audio to extract the features. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***005***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Bidirectional-GRU shall process the audio features step by step, making a prediction for each frame and output probability matrix for characters. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***006***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The CTC decoder shall take this probability matrix and extract the highest probability characters . | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SR\_***008***-V1.0 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The model shall return the output to ECU. | | |