## Customer Requirements Speciﬁcation

## **SAG ML Speech\_Recog CRS**

ـــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــ

**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\_ Speech\_Recog | |
| Version | V1.2 | |
| Status | Proposed | |
| Author | AES | |
| Date | [18-11-2022] | |
| Team approval | AA | approved |
| Mentor approval |  |  |
| Final approval |  |  |

# 

# Document History:

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Author** | **Date** | **Change** |
| 1.0 | AES | [10-11-2022] | * Initial Creation * Add CRS requirements |
| 1.1 | AA | [13-11-2022] | * Edit definition * Edit title of folder * Edit key elements |
| 1.2 | AES | [18-11-2022] | * Edit the sections alignment * Edit the document status. * Edit the project Definition. * Edit the document name. * Edit the system context * Edit all the crs requirements except the last one. * Add inner signals information. |

# Reference Document:

|  |  |  |  |
| --- | --- | --- | --- |
| **Ref.number** | **Doc.Name** | **Version** | **Status** |
| 1 | PO\_SAG\_CR\_ML | V1.2 | Released |

# Project Description:

## Definition:

The speech recognition model is supposed to receive the voice command of the user in the form of an audio file and convert it to text that will correspond later to the action the user wants to be performed.

## Features:

* Speech recognition

## Key Elements:

* The speech recognition model shall recognize user’s voice.

## **Speech Recognition:**

## System Context:

## 

## CRS Requirements:

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeechRecog\_***001***-V1.2 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recognition\_Class shall transform the Input\_Audio\_Files into INN\_Mel\_Spectrograms | | |

|  |  |
| --- | --- |
| **Info\_ID** | INFO\_PO\_SAG\_CRS\_ML\_ SpeechRecog\_***001***-V1.0 |
| **Description** | The INN\_Mel\_Spectrogram is an inner signal which is processed through the speech recognition process. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeechRecog\_***002***-V1.2 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recognition\_Class shall encode character label for each audio file into some numerical value. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeechRecog\_***003***-V1.2 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recognition\_Class shall use ResCNN to process the INN\_Mel\_Spectrograms and extract the INN\_AudioFeatureMap | | |

|  |  |
| --- | --- |
| **Info\_ID** | INFO\_PO\_SAG\_CRS\_ML\_ SpeechRecog\_***002***-V1.0 |
| **Description** | The INN\_AudioFeatureMap is an inner signal from the ResCNN and it represents the feature maps for the audio which are processed through the speech recognition process |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeechRecog\_***004***-V1.2 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recognition\_Class class shall use Bidirectional-GRU to take the INN\_AudioFeatureMap making a prediction for each frame and output the INN\_SoftMaxProbablities. | | |

|  |  |
| --- | --- |
| **Info\_ID** | INFO\_PO\_SAG\_CRS\_ML\_ SpeechRecog\_***003***-V1.0 |
| **Description** | The INN\_SoftMaxProbablities is an inner signal that the bidirectional- GRU outputs and it represents the SoftMax probabilities over the vocabulary. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeechRecog\_***005***-V1.2 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recognition\_Class shall use the CTC decoder to get  the INN\_SoftMaxProbablities and extract the highest probability  characters. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeecRecog\_***006***-V1.2 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recognition\_Class class shall return the Output\_Text\_Command to ECU. | | |