## 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 | V1.3 | |
| Status | Proposed | |
| Author | EID | |
| Date | [19-11-2022] | |
| Team approval | SAM | approved |
| Mentor approval |  |  |
| Final approval |  |  |

# 

# Document History:

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Author** | **Date** | **Change** |
| 1.0 | EID | [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 | EID | [18-11-2022] | * CRS\_REVIEW\_001 * CRS\_REVIEW\_002 * CRS\_REVIEW\_003 * CRS\_REVIEW\_005 * CRS\_REVIEW\_006 * CRS\_REVIEW\_007 * CRS\_REVIEW\_008 * CRS\_REVIEW\_009 * CRS\_REVIEW\_0010 |
| 1.3 | EID | [19-11-2022] | * CRS\_REVIEW\_004 * CRS\_REVIEW\_0011 * CRS\_REVIEW\_0012 |

# Reference Document:

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

# Project Description:

## Definition:

Speech Recognition is the process of converting human sound signals into words or instructions.

The goal of speech recognition is to transform the input speech feature vector sequence into a sequence of words using phonetic and linguistic information.

## 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***-V1.3 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recog\_Class shall transform the Input\_Audio\_Record 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***-V1.3 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recog\_Class shall encode character label for each audio file into some numerical value. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Req\_ID** | Req\_PO\_SAG\_CRS\_ML\_SpeechRecog\_***003***-V1.3 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recog\_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***-V1.3 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recog\_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***-V1.3 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recog\_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***-V1.3 | **Covers** | PO\_SAG\_CR\_ML\_006-V1.2 |
| **Description** | The Speech\_Recog\_Class shall return the Output\_Text\_Command to ECU. | | |