



Australian Government

ICTAI503 Build natural language processing models and pipelines

Release: 1

ICTAI503 Build natural language processing models and pipelines

Modification History

| Release | Comments |
|-----------|---|
| Release 1 | This version first released with the Information and Communications Technology Training Package Version 8.0. Newly created unit of competency to address in-demand skills needs. |

Application

This unit describes the skills and knowledge required to develop and apply routine natural language processing (NLP) pipelines to support the configuration of organisational computers to decode, interpret and manipulate human speech and text.

The unit applies to individuals who work across a wide range of information and communications technology (ICT) roles, including support technicians, system administrators, programmers and cloud computing engineers.

No licensing, legislative or certification requirements apply to this unit at the time of publication.

Unit Sector

Artificial intelligence

Elements and Performance Criteria

| ELEMENT | PERFORMANCE CRITERIA |
|--|---|
| <i>Elements describe the essential outcomes.</i> | <i>Performance criteria describe the performance needed to demonstrate achievement of the element.</i> |
| 1. Prepare to create NLP process | 1.1 Confirm work brief and tasks according to organisational policies and procedures 1.2 Establish requirements for morphological, lexical, syntactic and semantic analysis, disclosure integration and pragmatic analysis 1.3 Obtain human language data sources according to work brief |

| ELEMENT | PERFORMANCE CRITERIA |
|---|--|
| 2. Extract, model and predict language tokens | 2.1 Identify fields and blocks of language content 2.2 Source tokenizer according to organisational policies and procedures 2.3 Identify and mark sentence, phrase and paragraph boundaries 2.4 Normalise and tag acronyms in document according to work brief 2.5 Predict parts of speech for each generated token |
| 3. Process texts and pipeline tokens | 3.1 Lemmatise text in required document according to work brief 3.2 Identify and filter out stop words in document 3.3 Assign grammatical components with dependency parsing tags 3.4 Define relationships between parent words and other words in required document 3.5 Identify noun phrases according to work brief 3.6 Conduct named entity recognition (NER) and label nouns with real-world concepts 3.7 Create pipeline and run each token through NER tagging model 3.8 Run coreference algorithm on each token |
| 4. Finalise NLP pipeline and required documentation | 4.1 Review and apply security classifications and required security level in consultation with required personnel 4.2 Secure and save NLP pipeline according to organisational policies and procedures 4.3 Seek and implement feedback on processes from required personnel |

Foundation Skills

This section describes those language, literacy, numeracy and employment skills that are essential to performance but not explicit in the performance criteria.

| Skill | Description |
|----------|--|
| Numeracy | <ul style="list-style-type: none"> Extracts, interprets and comprehends statistical data in the context of developing NLP pipelines Calculates data inputs and outputs to interpret metric results and outputs and evaluate accuracy |

| Skill | Description |
|---------------------------|---|
| Reading | <ul style="list-style-type: none">• Interprets technical information presented in graphic, diagrammatic and visual form |
| Writing | <ul style="list-style-type: none">• Compiles reports with input from a range of sources |
| Problem solving | <ul style="list-style-type: none">• Applies problem-solving processes to identify actions required to support resolution of organisational problems |
| Initiative and enterprise | <ul style="list-style-type: none">• Identifies and follows organisational protocols and procedures |

Unit Mapping Information

No equivalent unit. Newly created unit.

Links

Companion Volume Implementation Guide is found on VETNet - -

<https://vetnet.gov.au/Pages/TrainingDocs.aspx?q=a53af4e4-b400-484e-b778-71c9e9d6aff2>