Compiled by Dr T Oberholster
AI-900 version: October 2021

aRtVerse
ART + SCIENCE + NATURE

Website & Blog
aRtVerse Shop
Amazon Books
Teepublic Shirts
SpreadShirt Apparel

ML Technique mines insights from unstructured text

- Understands written and spoken language
- Analyse & interpret text with AI
- Interpret spoken language & synthesise speech response
- Translate spoken or written language
- Interpret commands and take actions
- Evaluate different aspects of document/phrase to gain insights into the context of the text.

Concepts

Build an ML Model

Step 1: Provision a Resource

- Text Analytics/Translator Text/Speech/LUIS resource: Dedicated training/ prediction or both resources. Use this resource only, or manage access/ utilization/ tracking/ billing of resource separately. Cannot use authentication key/ service endpoint interchangeably.
- Cognitive Services (CS) resource: Dedicated training/ prediction or both resources. Combine Custom Vision or other services (Text Analytics, Translator Text), or manage access/ utilization/ tracking/ billing of resource together. One endpoint and authentication key for resources.
- Provide client application authenticaion key and HTTP address endpoint for ML service consumption

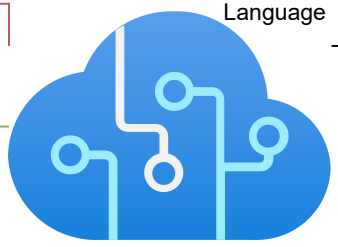
MS AZ Computer Vision

- Subset of Analyze Image of Computer Vision
- Pretrained NLP ML models and services
  - **Text Analytics:** Analyse text documents and extract phrases, detect entities (places, dates, people), and evaluate sentiment (positive/negative)
  - **Translator Text:** Translate +60 languages
  - **Speech:** Recognize, synthesise, and translate speech
  - **Language Understanding Intelligent Service (LUIS):** Train language model to understand spoken or text-based commands

Services

Cognitive Services for Language

Speech



Natural Language Processing & Conversational AI

Use Cases

- ID and categorize concepts
- ID main points in unstructured text
- Understand customer perception
- Process unstructed medical/clincal data
- Conversational interface with FAQs/manuals/website
- Automate workflow by classifying unstrucuted text

Text

Speech service

- Speech-to-text
- Text-to-speech
- Speech translation
- Speaker recognition

- AI accepts vocal commands and returns spoken responses (home and auto systems)
  - **Speech recognition:** Ability to detect & interpret spoken language
  - **Speech synthesis:** Ability to generate spoken output

Text Analytics Service

- Statistical analysis of terms (remove stop words) with frequency analyses (counting words) for clues on main topic
- Extend frequency analyses to multi-term phrases
  - N-grams, a two-word phrase = bigram and three-word phrase = trigram
- Stemming or lemmatization algorithms to normalize words before counting (only use root words)
- Analyse sentences with linguistic structure rules (noun, verbs, and adjectives)
- Encode words/terms as numeric features to train ML model.
  - Classify document based on terms inside.
  - Sentiment analyses to classify document as positive/negative
  - Create vectorized models to capture semantic relationships between words in n-dimensional space. Related terms are grouped closer together.

Speech recognition (Speech service)

- Transcribes spoken word (recorded or live) into data for processing
- Speech patterns analysed & mapped to words using multiple models
  - **Acoustic model:** Converts audio signal into phonemes (represent specific sounds)
  - **Language model:** Maps phonemes to words (using statistical model and confidence scores)
- Words converted to text
  - Closed captions for videos
  - Transcript of phone call or meeting
  - Automated note dictation
  - Determine intent of user for processing

Text Analytics Service

- **Language detection:**
  - Language name
  - ISO 6391 language code
  - Confidence score of detection
  - Focus on predominant language in the text in mixed language submissions
  - Ambiguous language (keyboard emoji) evaluated as unknown with NaN score
- **Sentiment analysis:**
  - Evaluate & return sentiment scores & labels for sentences
  - Range 0 – 1, 1 = positive, 0.5 = neural/indeterminate, 0 = negative
  - Indeterminate may be due to insufficient context, no sentence structure or incorrect language used (= 0.5)
- **Key phrase extraction:**
  - Evaluate text & identify main topic(s)
  - Entity recognition
  - List of entities (people, places, dates, objects)
    - Categories (person, location, organization, quantity, datetime, URL, email, phone number, IP address)
    - Subtypes (number, percentage, ordinal, age, currency, dimension, temperature, date, time, daterange, timerange, duration or set)
  - Entity linking to disambiguate entities using a specific reference.

Speech synthesis (Speech service)

- 'Reverse' of speech recognition
- Vocalising data, text converted to speech
  - Text to be spoken
  - Voice used to vocalize speech
- Systems tokenizes text (broken down into word) & assigns phonetic sounds to words
  - Phonetic transcription broken down into prosodic units (phrases, clauses, or sentences) to create phonemes to convert to audio and synthesised to words using a voice
  - Voice determines pitch and timbre in audio wave form to output to speaker or written to file
- Used for
  - Generating spoken responses to user input
  - Voice menus for telephone systems
  - Read email or text messages (hands free)
  - Broadcast announcements (public locations, stations, or airports)

Speech-to-Text API (Speech service)

- Convert text to audible speech, played through computer speaker or saved to file
- Pre-built voices with multi-language & regional pronunciation support
- Standard and neural voices (leverage neural networks to create intonation and more natural sounding voices)
- Develop custom voices

Translate Text and Speech

- Automated or machine translation allows people from different culture and/or geographical location to collaborate
- Multi-language support for text translation, transliteration, language detection, and dictionaries.
- **Literal translation:** Each word translated to corresponding word. Issues include no equivalent words or changes in phrase/meaning/context.
- **Semantic translation:** AI understands context of the words used and returns a more accurate translation where grammar, formal/informal and colloquialisms are considered.
- **Text translation:** Translate documents from one language to another (email, web pages)
- **Speech translation:** Translate between spoken languages, directly (speech-to-speech) and via text (speech-to-text-translation).

Translator Text Service and Speech Service

- For applications, websites, or tools
  - Neural Machine Translation (NMT) analyses semantic context of text and renders more accurate/complete translation
  - 60+ languages support
  - Provide ISO 639-1 language or 3166-1 cultural codes for languages being translated
  - One-to-one and one-to-many translations

- Optional configurations
  - **Profanity filtering:** Culture specific profanity removed by marking or omitting results
  - **Selective translation:** Tag content not to be translated (code, brand, name, word/phrase)
- **Speech service:** speech-to-text and speech-to-speech translation
  - Speech-to-text, Text-to-speech, and Speech translation
  - **Speech translation:** translate spoken word from streaming audio (live or recorded) for real-time closed captioning or two-way communication
    - One-to-one or One-to-many translations
    - Specify language and/or culture code format to consider localized pronunciation and linguistic idioms

Language Understanding

- MS Azure Language Understanding Intelligent Service (LUIS)
  - **Utterances:** Question/phrase/command
    - Different ways of saying the same thing
    - Used to train the model
  - **Entities:** An object/person/place to which is being referred
    - To which an action is applied by the model
    - List: Entities in hierarchical list/sublist and includes synonyms
    - RegEx: Designed as regular expression (based on a pattern of alphanumerical characters)
    - Pattern.any: Patterns for complex entities that may be hard to extract from sample utterances
  - **Intents:** Purpose/goal as determined by the utterance
    - Action for model to execute on an entity
    - There are numerous utterances, grouped by intent and related to entities
    - None intent is a fallback for unknown utterances with generic responses
- After model training (authoring), the model predicts the intent and entity based on the utterance to respond appropriately (prediction).
- Two resources made for Authoring and Prediction to separate tracking of resource utilization
  - Authoring
    - Author and train LU application with utterances, intents & entities
    - Prebuilt domains with predefined intents and entities for common scenarios
    - Custom intents and entities and map to utterances
    - Create model through web-based LU portal or with code
    - After training responses can be reviewed and updated.
  - Predicting
    - Trained LU model is published to application for consumption
    - Connect to a client application endpoint, specify authentication key and submit user input for predictions and return an action based on intent

Conversational AI

- Bots manage and participate in conversations with humans
- Dialogs through website, email, social media, messaging, phone calls and other channels
- Customer product/services, reservation systems, health care consultants, home automation

QnA Maker Service

- Build a knowledge base of question-and-answer pairs, usually with built-in NLP for multiple phrased examples but same semantic meaning.
- Code with REST API or SDK or use QnA Marker portal
- QnA pairs created from existing FAQ or page, imported from chit-chat data source, or manually created or combo
    - Alternative phrasing consolidates questions with same meaning
    - Train knowledge base model with literal QnA then test with built-in interface and review answers
- Publish to client applications with knowledge base ID, endpoint, and authorization key
    - Deliver knowledge base via a bot, created in MS Bot Framework SDK to write conversational flow and integration code
    - Or use automatic bot creation in QnA Maker via Azure Bot Service
- Manage the bot in Azure portal
    - Extend functionality with code (simple edit in Azure portal, comprehensive ~ download and edit locally and republish to Azure)
    - Test the bot's interactive interface
    - Config logging, analytics, and integration with other services
    - Connect bot to multiple channels, users interact through chat, email, Teams, and other media.

Azure Bot Service

- Develop, publish, and manage bots.
- Bot Framework to create bot, manage with Bot Service and integrate QnA/LUIS for interactions.

Challenges & risks

- **Bias:** Race or gender bias in model due to bias in data
- **Errors:** When system fails, drastic consequences
- **Breach:** Sensitive data leaked/stored insecurely
- **Suboptimal solutions:** Caters to specific user, not universal
- **Trust:** Users do not know about underlying data/assumptions
- **Liability:** Who is responsible when predictions are wrong