**Artificially Intelligent Digital Assistant (AIDA)**



**Pavly Salah Zaki – section 2**

**Bolis Karam Soliman – section 2**

**Marco Magdy William – section 3**

1. **Introduction**

Artificially Intelligent Digital Assistant (AIDA) is an Artificial Intelligence**[1]** and Speech Recognition**[2]**-powered digital assistant. AIDA is designed to help the impaired communicate more easily with their devices and do complicated tasks just by saying commands. It is mainly designed for desktop devices at the moment, but support for various operating systems is soon to come.

1. **Speech Recognition**
2. What is Speech Recognition?

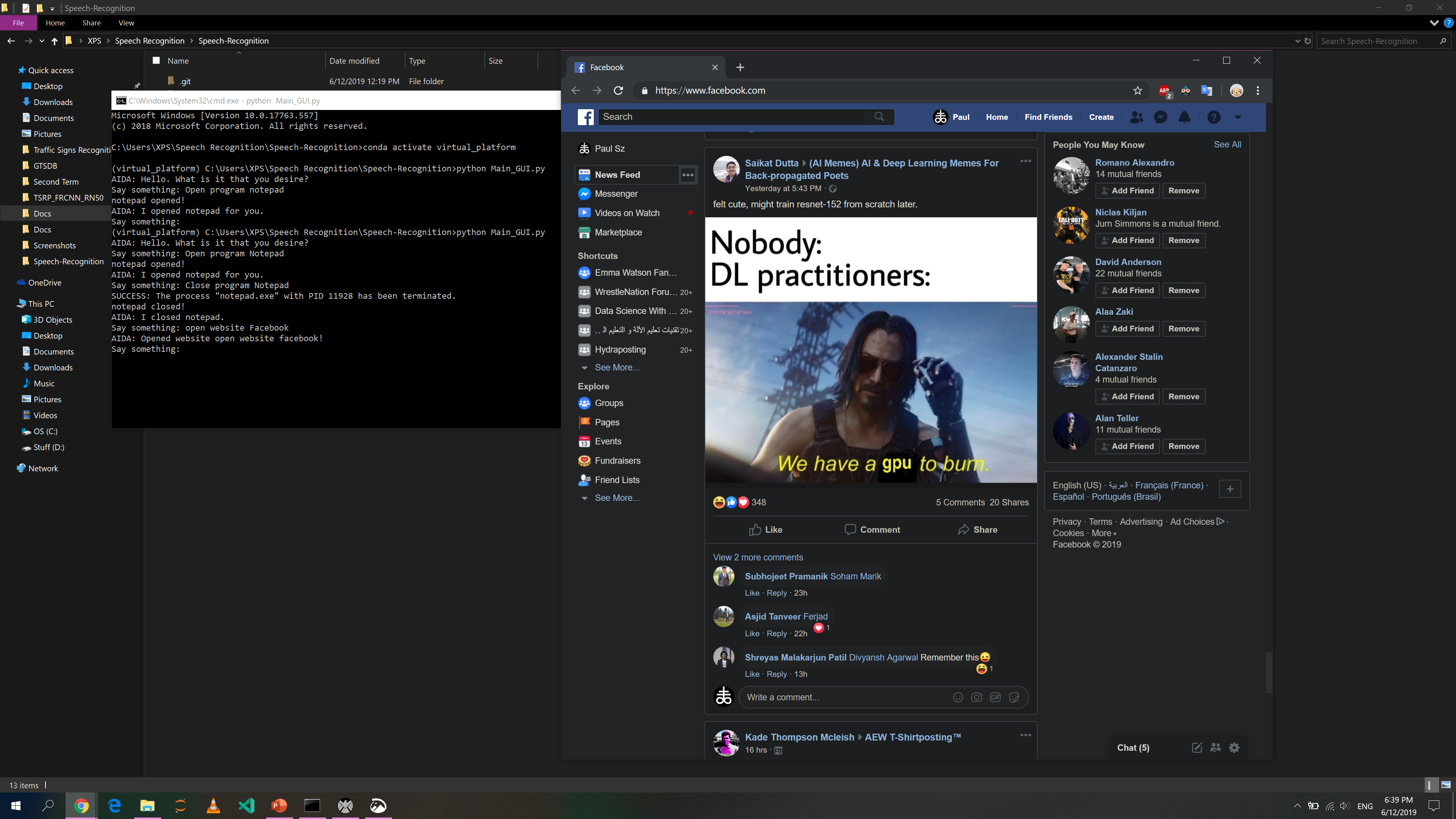
Speech Recognition is the technology that allows humans to interact with machines through the aid of vocal commands. Speech Recognition allows machines to understand the words uttered by the speaking human and execute their commands.

1. How does Speech Recognition Work?
   1. Pre-processing

* When a person speaks, the vibrations of the vocal cords produce a wave. This wave is the audio/sound signal.
* The signal is passed through a Noise Cancellation or Reduction phase in order to eliminate as much noise as possible while maintaining as much original information as possible.
* The audio signal is an analog signal(continuous time and continuous amplitude), but a computer can’t understand such a signal, therefore it must be converted into a digital signal (continuous time and discrete amplitude) using a sampler or an Analog-to-Digital Converter (ADC) **[3]**. Even then, the continuous time aspect would be an incredibly tiresome for a computer to handle, hence the use of a quantizer to make the time also discrete**[4]**.
  1. Digital Signal Processing
     1. The signal is segmented for as few as 10ms. This way we can treat the signal as a series of constant, stationary values, each value may represent use information.
     2. Each segmented signal is then passed through a Hidden Markov Model**[5]**which produces a 10-dimensional vector output.
     3. Each vector is then mapped to one or more phonemeusing a Speech Decoder.
     4. A special Artificial Neural Network**[6]** then groups phonemes together to form a comprehensible word.

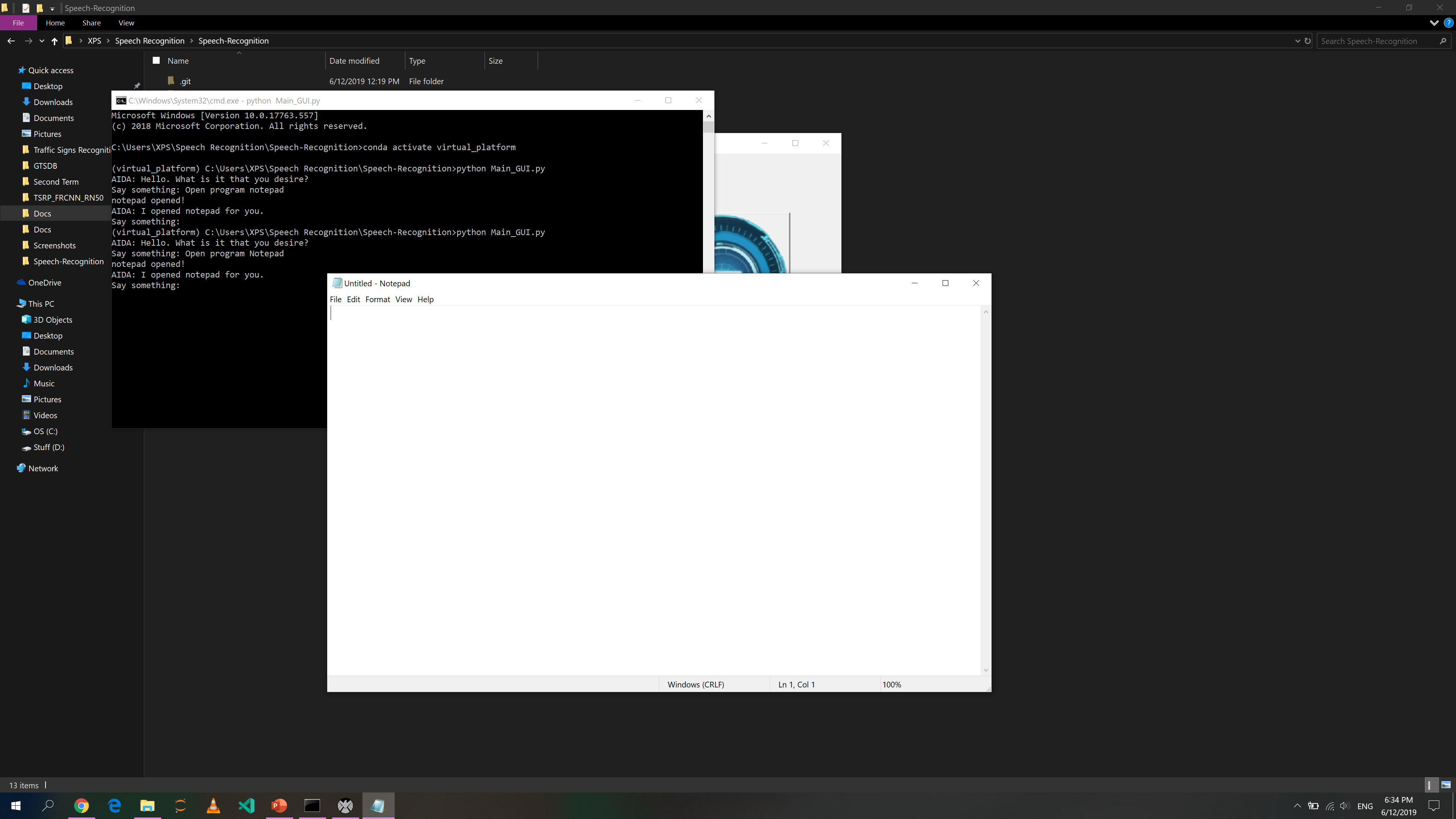
1. **Available Commands**
2. **Open a web site**

* *Use:* Opens a particular website
* *Command:* “Open website {website name}”
* *Raises an error:*
  + No internet connection
  + Invalid website name
* *Example:*

**

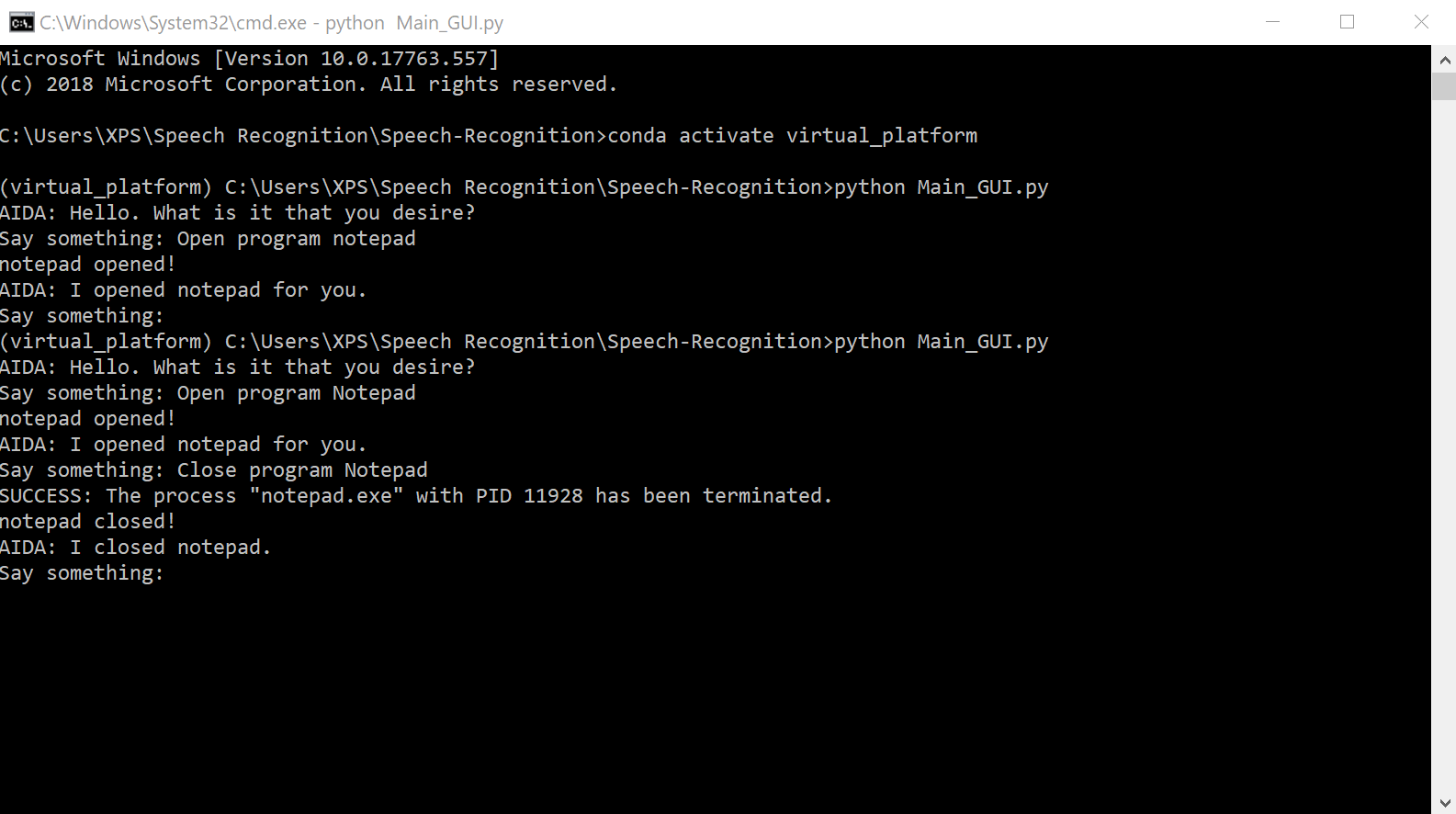
1. **Open a program**

* *Use:* Opens a particular program
* *Command:* “Open program {program name}”
* *Raises an error:* Program is not installed
* *Example:*

****

1. **Close a program**

* *Use:* Closes a previously opened program
* *Command:* “Close program {program name}”
* *Raises an error:* Program is not opened
* *Example:*

****

1. **Send an e-mail**

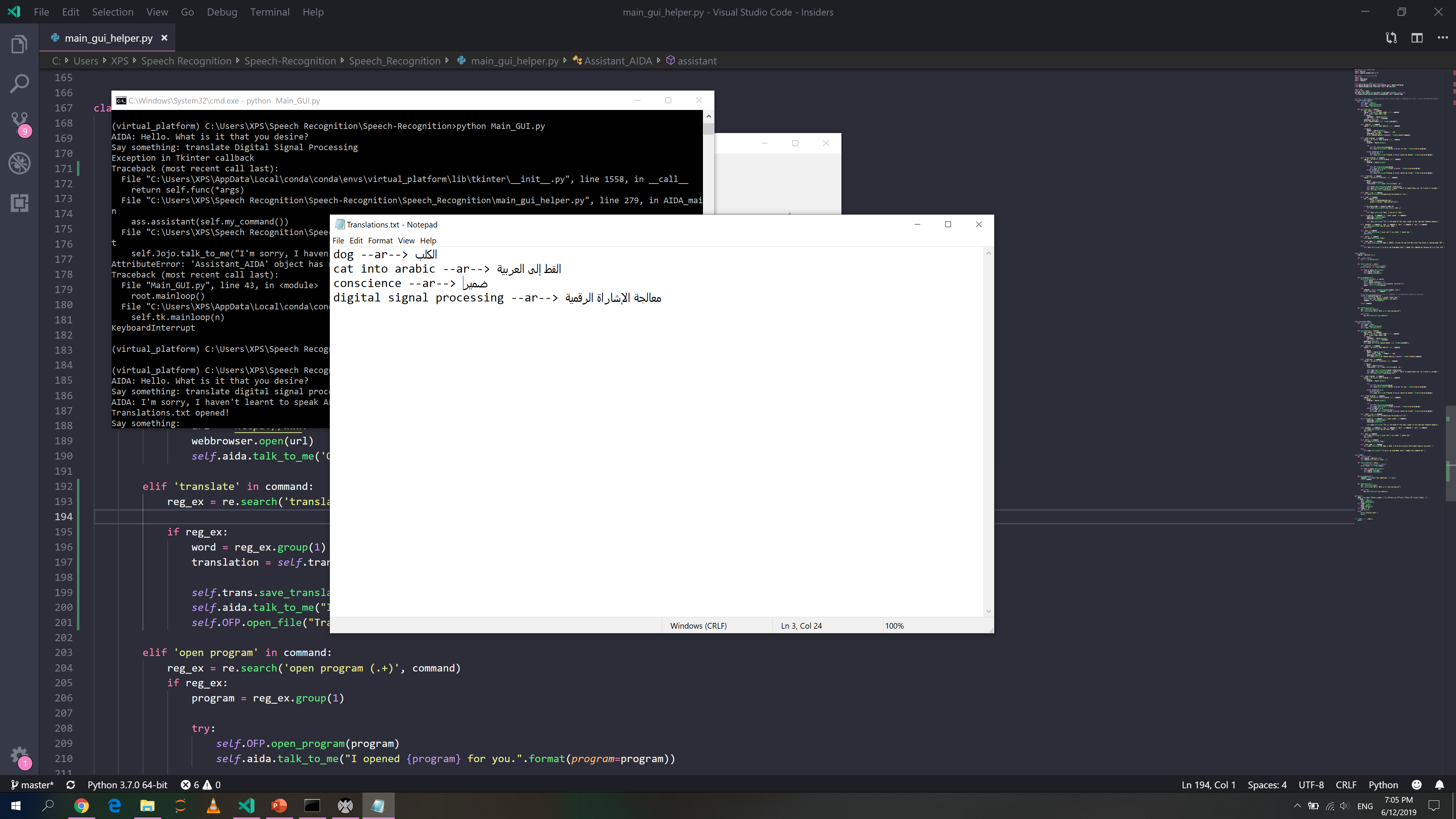
* *Use:* Send an e-mail to a person
* *Command:* “Send an e-mail to {person name}, and say {message}”
* *Raises an error:* 
  + - No internet connection
    - Person cannot be reached
* *Example:*

1. **Weather forecast**

* *Use:* Tells the weather forecast today
* *Command:* “Weather forecast”
* *Raises an error:* No internet connection
* *Example:*

1. **Translation**

* *Use:* Translate a word into a specific language
* *Command:* “Translate {word} into {language}”
* *Raises an error:*
  + 1. No internet language
    2. Invalid language
* *Example:*

****

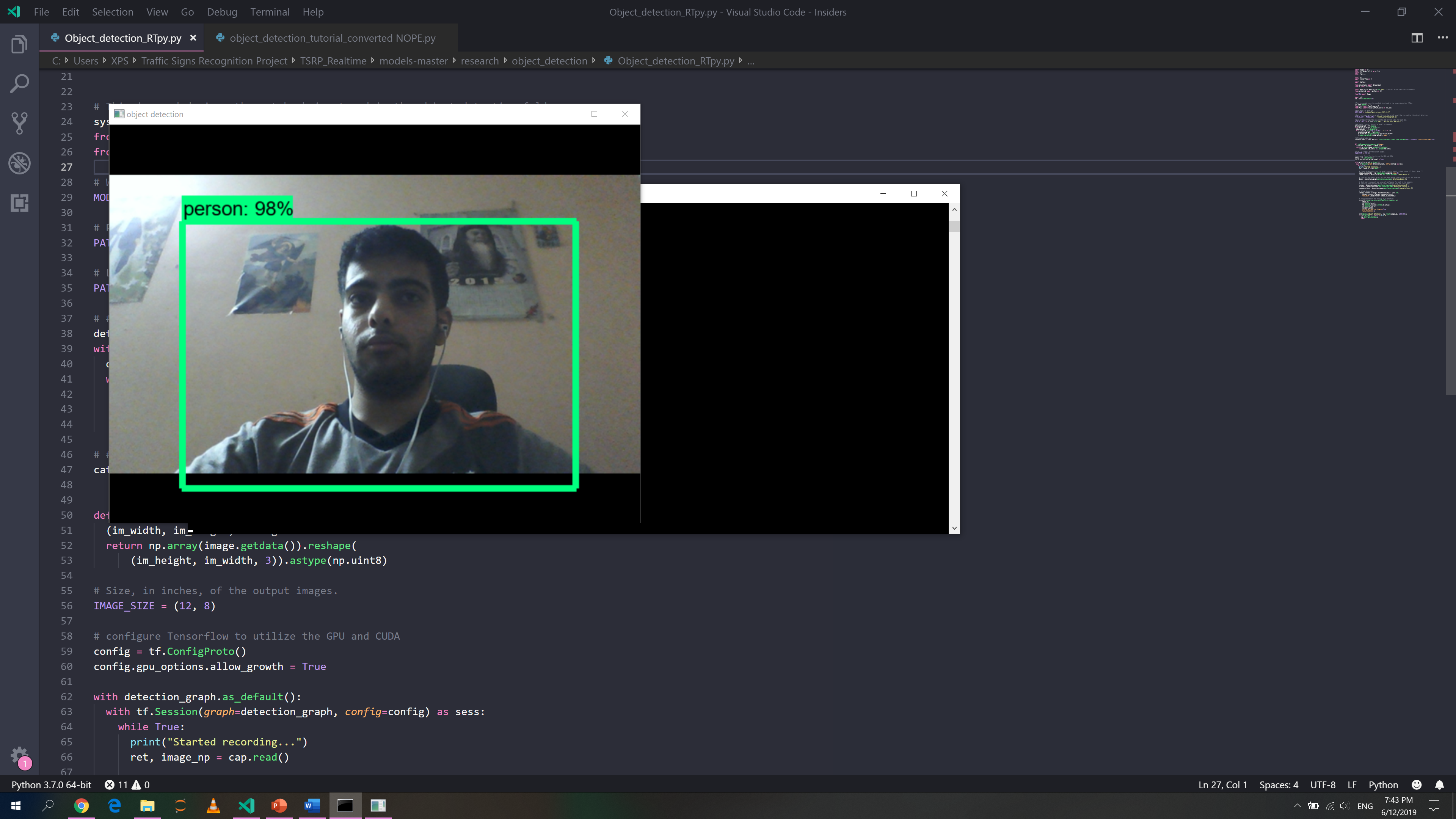
1. **Audio visualization**

* *Use:* Visualizes the input sound signal in the time and frequency domains
* *Command:* “Visualize input signal” or “Visualizer”
* *Raises an error:* None
* *Example:*

****

1. **Object detection**

* *Use:* Uses a deep learning model to detect objects around the user
* *Command:* “Object detection” or “Detector”
* *Raises an error:* None
* *Example:*

****

1. **Future Work**

* ***Mobile application***
* ***Performance Optimization***
* ***Commands***
  1. TODO notes
  2. Call {Person} (for mobile)
  3. Take a picture or a video
  4. Play a song
  5. Google search
  6. Sports results
  7. Calculator

1. **References**
2. <https://en.wikipedia.org/wiki/Artificial_intelligence>
3. Pahini A. Trivedi, “Introduction to Various Algorithms of Speech Recognition: Hidden Markov Model, Dynamic Time Warping and Artificial Neural Networks,” 2014
4. <https://www.analog.com/en/products/analog-to-digital-converters.html>
5. <http://rs-met.com/documents/tutorials/DigitalSignals.pdf>
6. <https://web.stanford.edu/~jurafsky/slp3/A.pdf>
7. <https://www.sciencedirect.com/topics/engineering/artificial-neural-network>