

How to make a digital assistant, smarter!

-Dhruv Apte(the-ethan-hunt)

P.S. I am not Tom Cruise

Who am I?

- Third year undergrad majoring in mechanical engg. from GCET, Gujarat.
- Worked in Python for most of the time.
- Some out-of-the-box projects, web development and a some machine learning and NLP on the side.
- <https://github.com/the-ethan-hunt> is the base from where I conduct my not-so-clandestine activities.
- Participated in KWOC'16 as a contributor; this time as a mentor

Introduction

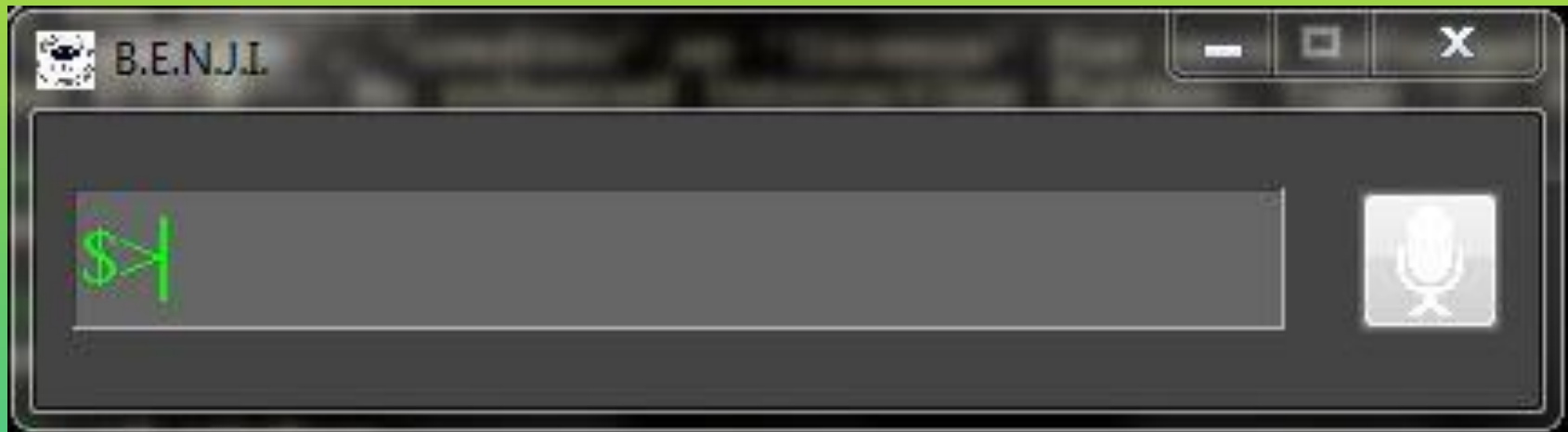
- In 2017, a guy named Mark made a personal project to recreate a assistant similar akin to the JARVIS of Iron-Man.
- He created it for his home a 'JARVIS' with the voice of another guy named 'Morgan'.
- And he could not open-source it, so he asked me to do it.

The start

- I decided to create my own digital assistant titled B.E.N.J.I. – Brilliant and Effective Non-Javascript-enabled Intelligence.
- It's function- Display me headlines of a press agency as I tell it by speech.
- 2 problems:
 - 1) It had a very very dull appearance.
 - 2) A very minor problem not worth mentioning here

KWoC'17

- I registered my project with some more features on KWoC; it got selected and work began.
- Everything was new:
 - 1) New GUI
 - 2) New logo
 - 3) New features
 - 4) New bugs



THE BIG QUESTION

- How can AI possibly help us out here?
- Can its other sub-parts like NLP, DL or ML help us out?
- Key definitions:
 - 1) NLP- natural language processing- the ability of a computer program to understand the language of humans.
 - 2) DL- deep learning – Feeding computer system a lot of data which it uses to make decisions about other data.
 - 3) ML- machine learning- provides the system to automatically learn and improve from exp. Without explicitly programmed.

- B.E.N.J.I. can help play a video on YouTube and has another option to download it.
- So, if I speak 'download Coldplay paradise' will it download the song or play any random song titled 'paradise' since it also contains paradise?
- How can B.E.N.J.I. recognize the position of 'play' or 'stream' is in the command or in the name of the song?
- What if I tell B.E.N.J.I. to play a Hindi or any regional language song? Can B.E.N.J.I. understand it?

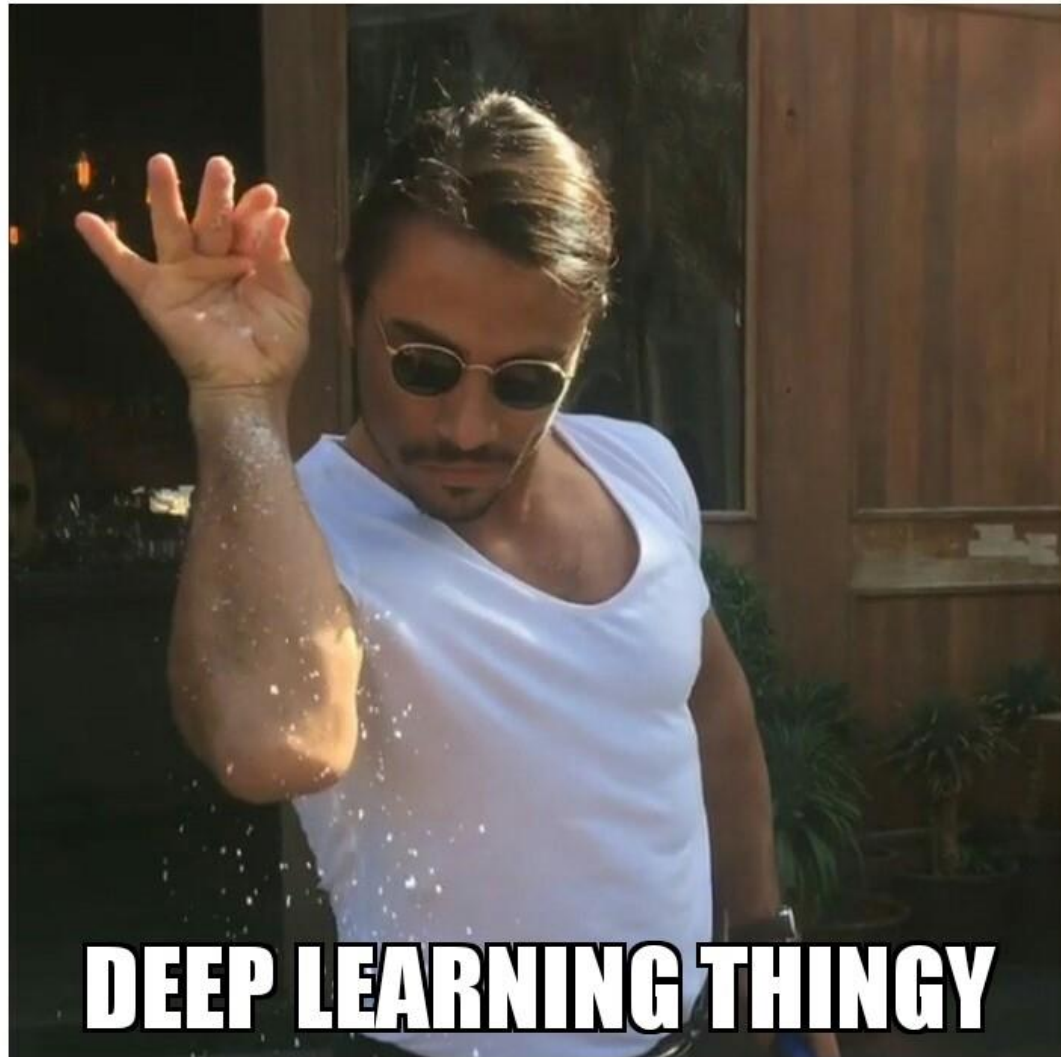
NATURAL LANGUAGE PROCESSING



Second Part of the problems

- What type of song I might want to listen comes to mind when I order B.E.N.J.I. to play a song.
- This happens when there are multiple songs with the same title. For example, a small search leads me to a music directory where I find there are 7 songs that are titled 'paradise'. There might be more.
- How can B.E.N.J.I. know what 'paradise' did I mean to listen?

LEMME SPRINKLE
SOME



DEEP LEARNING THINGY

More DL?

- Music Recommendations?
- https://github.com/oarriaga/face_classification- A face detection and emotion classification on a keras CNN model
- CNN- neural network that derives its name from a convolutional operator that extracts features from a given image.
- These network guys here are quite effective in recommender systems and image recognition.
- Movie recommendations?
- What about books?

Visual AI

- What if someone tries to unlock my device without my permission?
- We need visual AI to let the assistant recognize whether the person who's giving him commands is his true owner or not.
- Suggestion: camera at top of device with facial detection?

Conclusion

- Yes, AI can be beneficial, for sure! I have given examples of how it can help improve B.E.N.J.I.'s performance.
- But will be connected more and will deeply cling to the system; will have a lot of information about me too; probably making me more convinced not to open-source it.