SciPy India 2017

FAQ Bots for Freshmen:
*From AIML to Keras
with a stop at Chatterbot*

Saurabh Kumar

What are FAQ Bots?

- FAQ bots are a type of chatbots which are also known as an Artificial Conversational Agent.
- Simulate human conversation using a computer.
- Are like virtual assistants to whom you can ask questions or have a conversation.
- Retrieval based models / Generative models.

Why use them?

- Come at virtually zero cost.
- Can satisfactorily answer standard questions for which they have been trained on.
- Easy to distribute.
- No human intervention needed. Just build and deploy.
- Can learn from the new unseen queries like a human.

Examples Bots

FAQ Bots

- Livelier than dry FAQ pages.
- Availability.
- Offload human work of answering questions.
- Deploy as a mobile app or webpage.
- Computationally light training.
- Only important questions can be routed to the office staff.
- Can be linked to Text to Speech engines for more immersive experience.

AIML:

Artificial Intelligence

Markup Language

- Created by Richard Wallace ('95)
- XML style scripts for creating AI language agents.
- AIML interpreters - Python, Ruby, Java, C++…
- Open Source.
- Easy to use.
- ALICE uses AIML.
- AIML 2.0 - work in progress.

Elements of AIML

- **Categories:**
  - Basic unit of knowledge.
  - Contains a pattern and a Template and optional an context.
- **Patterns:**
  - String to be matched with one or more user inputs.
  - Supports wildcards.
- **Template**
  - Response to matched pattern.
  - Supports variables, redirection, conditional and random responses along with formatting options.

Examples

```
<category>
  <pattern>WHAT IS YOUR NAME</pattern>
  <template>My name is FAQ Bot.</template>
</category>



<category>
  <pattern>DID YOU KNOW *</pattern>
  <template>
    <random>
      <li>Be more specific.</li>
      <li>I didn't know that!</li>
      <li>Wow! That's interesting.</li>
    </random>
  </template>
</category>
```

- Machine Learning based conversational dialogue engine.
- Language independent - trained to converse in any language.
- FAQ bot for EE Dept here.
- Written in Python.
- Open Source.
- Learns on the go!

Elements of Chatterbot

- Simple YML style Training Corpus.
- Sample corpus provided with installation in 3 languages.
- Simple Python API.
- Adapts to the person the bot is communicating with in the long run.

Examples

```yaml
categories:

- conversations

conversations:

- - Good morning, how are you?

  - I am doing well, how about you?

  - I'm also good.
```

```yaml
categories:

- profile

conversations:

- - What are your interests

  - I am interested in all kinds of things. We can talk about anything!

- - What are your favorite subjects

  - My favorite subjects include robotics and natural language processing.
```
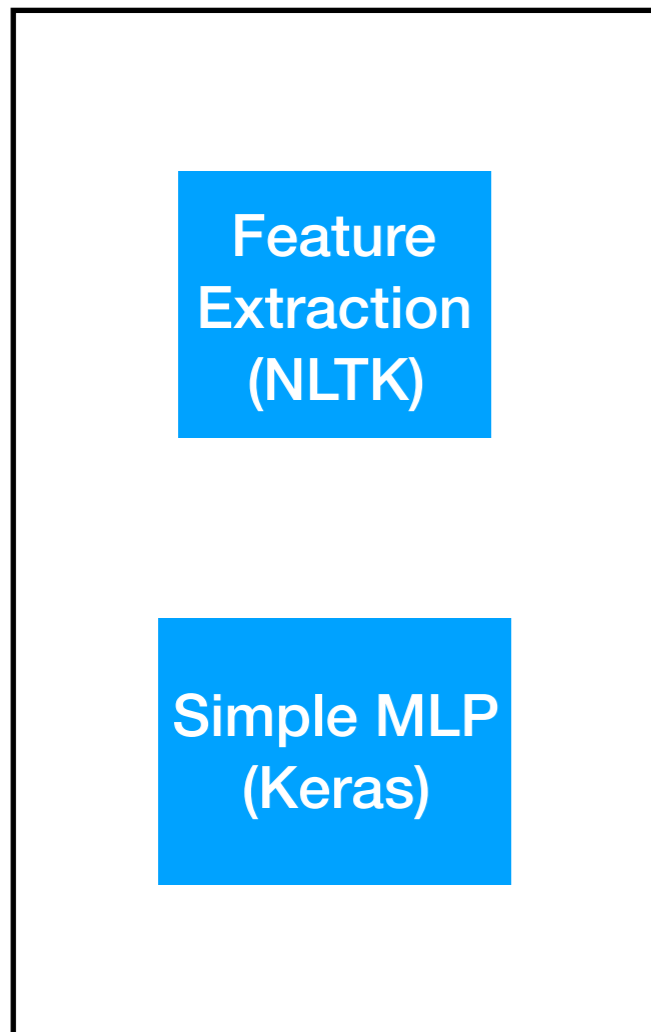
- High level neural networks API.
- Written in Python.
- Tensorflow/CNTK/Theano backend.
- Open Source.
- Easy to use.
- Designed for fast experimentation.
- Runs seamlessly on CPU/GPU.

Keras for FAQ bots

- Pose the conversation as a classification problem.
- Parse the user query with Natural Language Toolkit to construct features.
- Use these features to classify each query into a intent class.
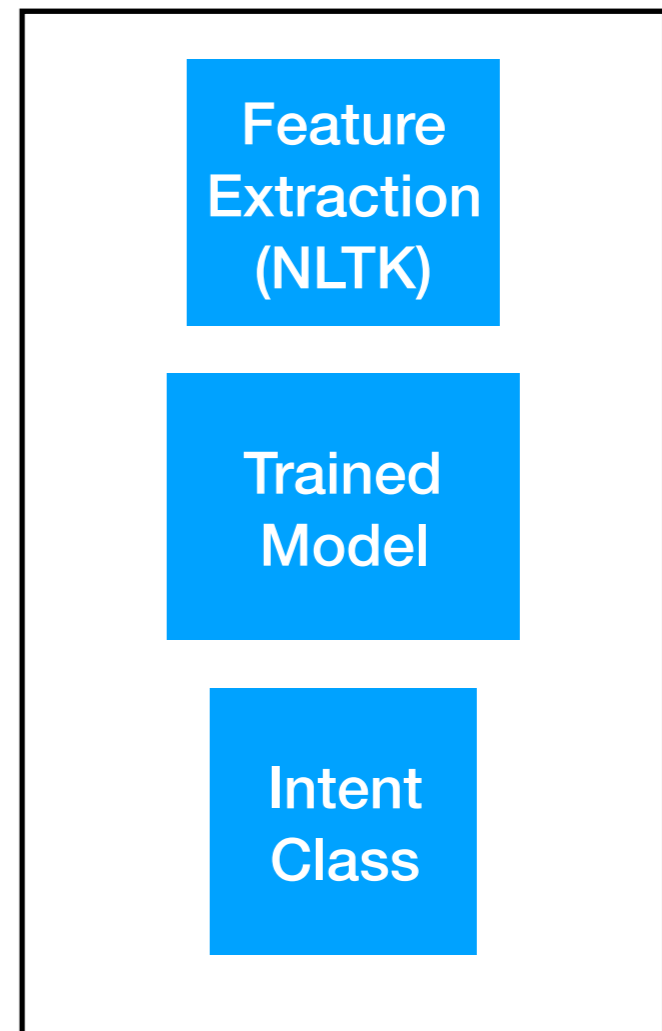- Simple JSON style specification of training corpus.

Training the bot

Labelled training corpus

Feature Extraction (NLTK)

Simple MLP (Keras)

Trained Model

Testing the bot

User Query (Flask)

Feature Extraction (NLTK)

Trained Model

Intent Class

Response for intent (Flask)

Further References and Tutorials:

- AIML

  - https://www.pandorabots.com/pandora/pics/wallaceaimltutorial.html
  - http://www.alicebot.org/documentation/aiml-primer.html

- Chatterbot

  - http://chatterbot.readthedocs.io/en/stable/tutorial.html

- Keras

  - https://keras.io
  - https://github.com/fchollet/keras-resources

Thank You!