Machine Learning Algorithms for Chatbots

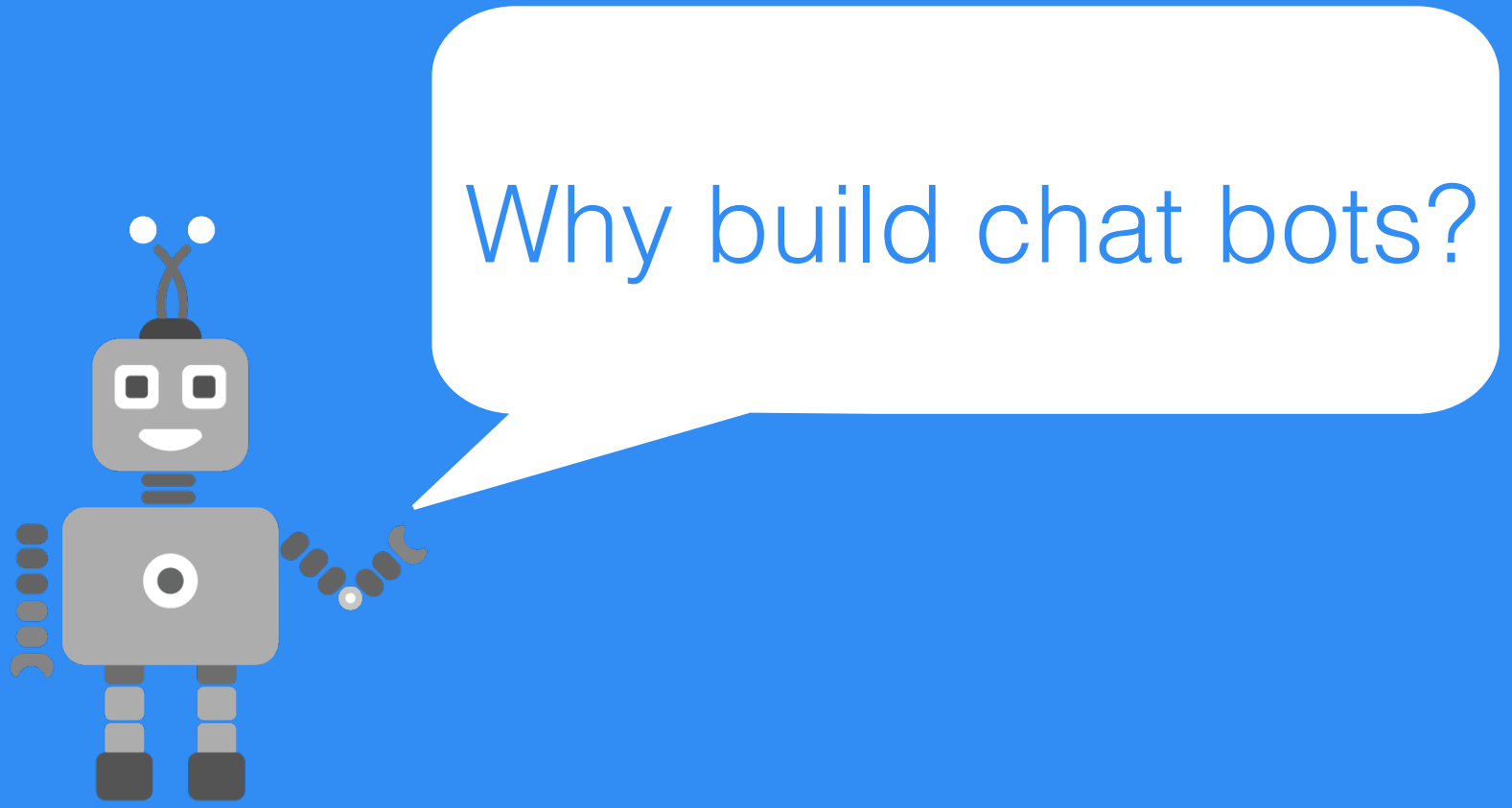
By

Krupal Modi

Machine Learning Scientist, Haptik

What are chat bots?

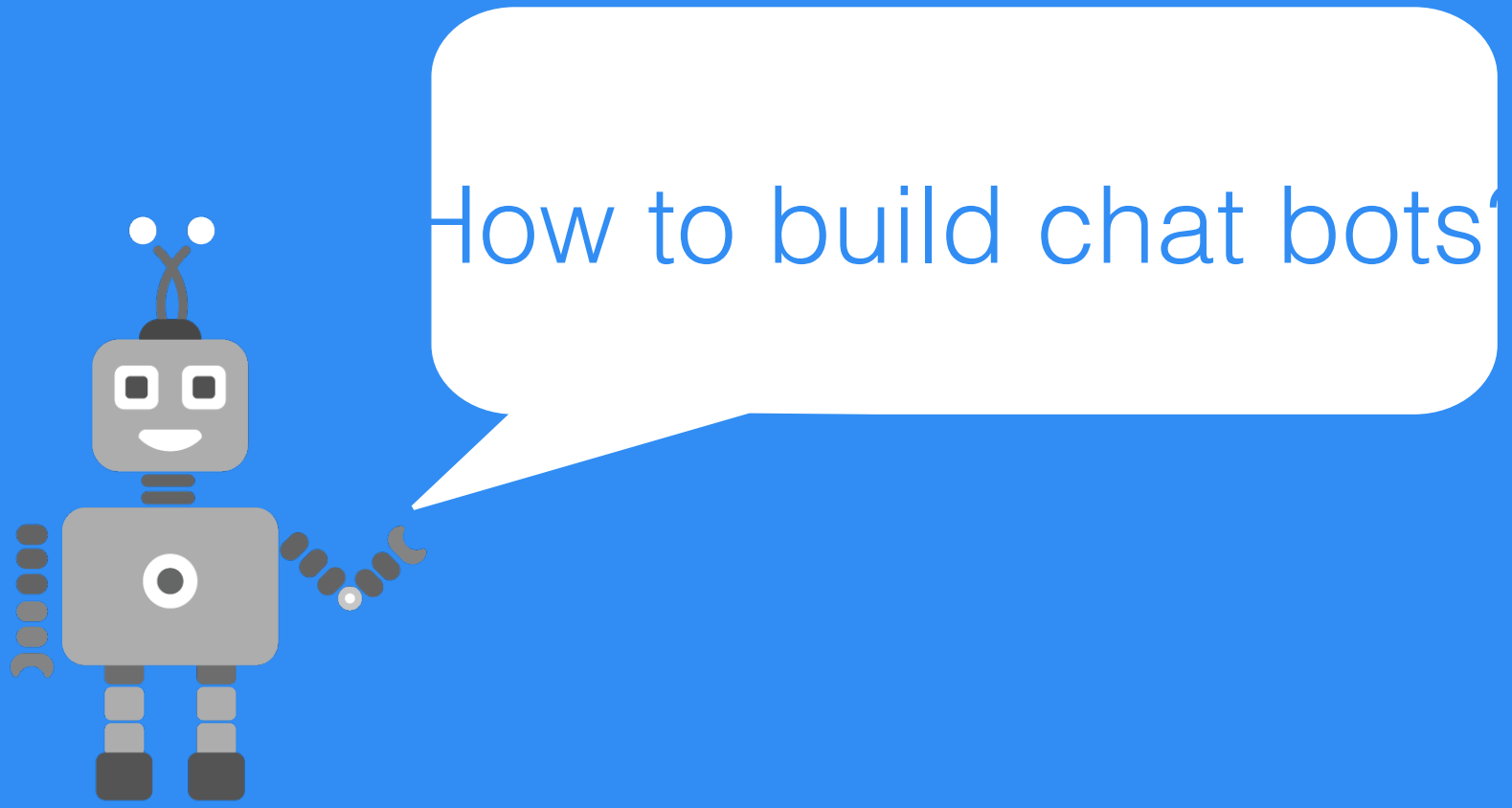
- Interface that enables users to interact with a machine through a conversational UI.
- Machine's need to understand the user's message and the context of the conversation to get things done.
- The interface of how people interact with the machine is as important as the algorithm.
  Design + Machine Learning.


Why build chat bots?

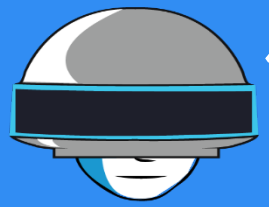
- Ease of doing everything on one screen
- Chat is the most commonly used interface on a smart phone
- Technology that feels like a friend

Application of Chat bots

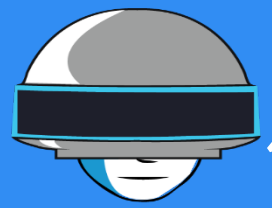
- Personal Assistant
- Health and nutrition
- Dating and love
- Education
- Customer care


How to build chat bots?

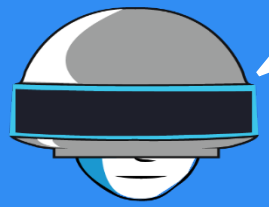
Too many decisions ?



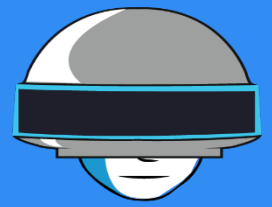
Wit.ai, Api.ai, SciPy
Scikit learn ?



TensorFlow, Theano,
Keras, RNNLM, CNTK ?



Facebook, slack,
Telegram, custom
App?



IBM Watson, AWS ML,
Microsoft Azure?

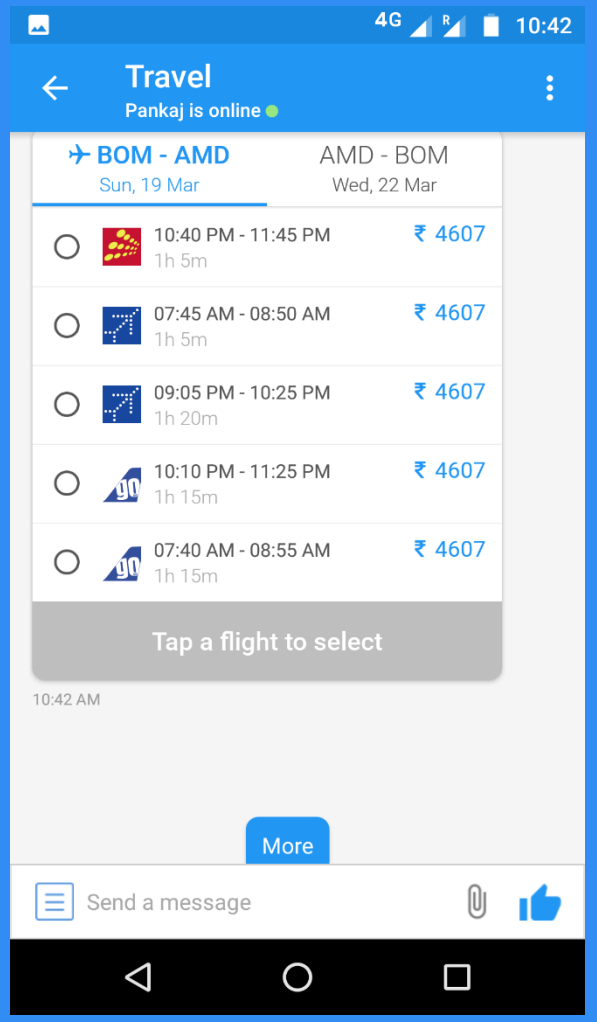
Getting Basics Right while building Chatbot

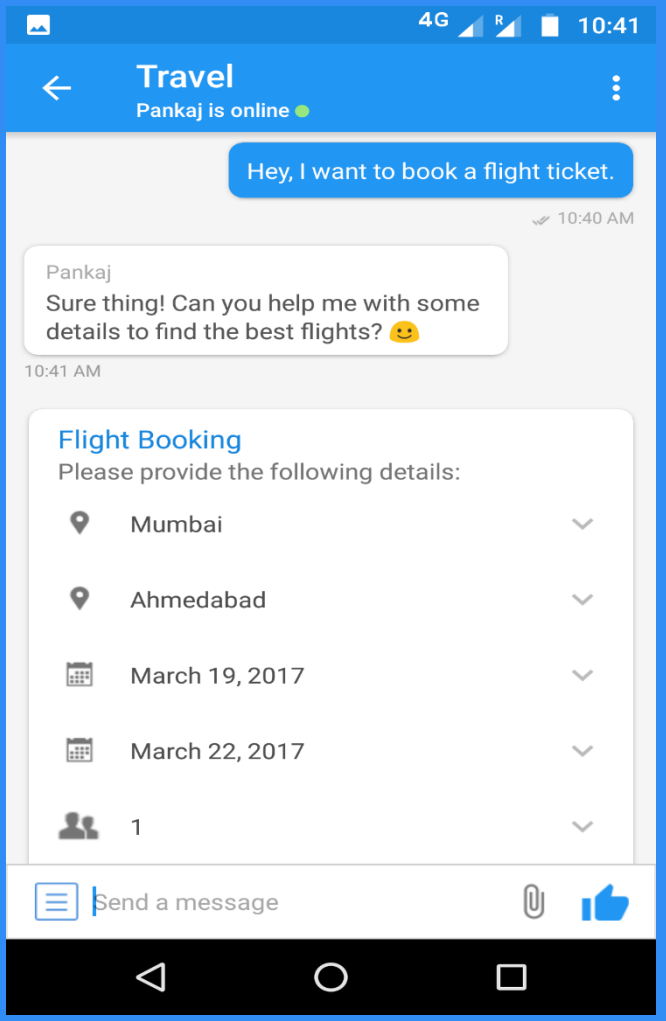
Collecting Training Data

- Volume and Accuracy of Training Data
- Domain of Data – Open Domain vs. Closed Domain
- Example Text :
- 1) User - “Go ahead with the booking” – Movies, Flights, Trains?
- 2) User – “Only Non stop flights”, Bot – “XYZ flight”, User – “Go ahead with the booking” - Flights

Preprocessing

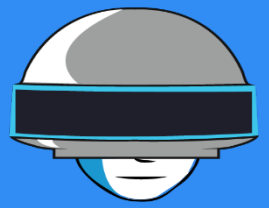
- Abstraction
- Latency
- Consistency





Feature Engineering

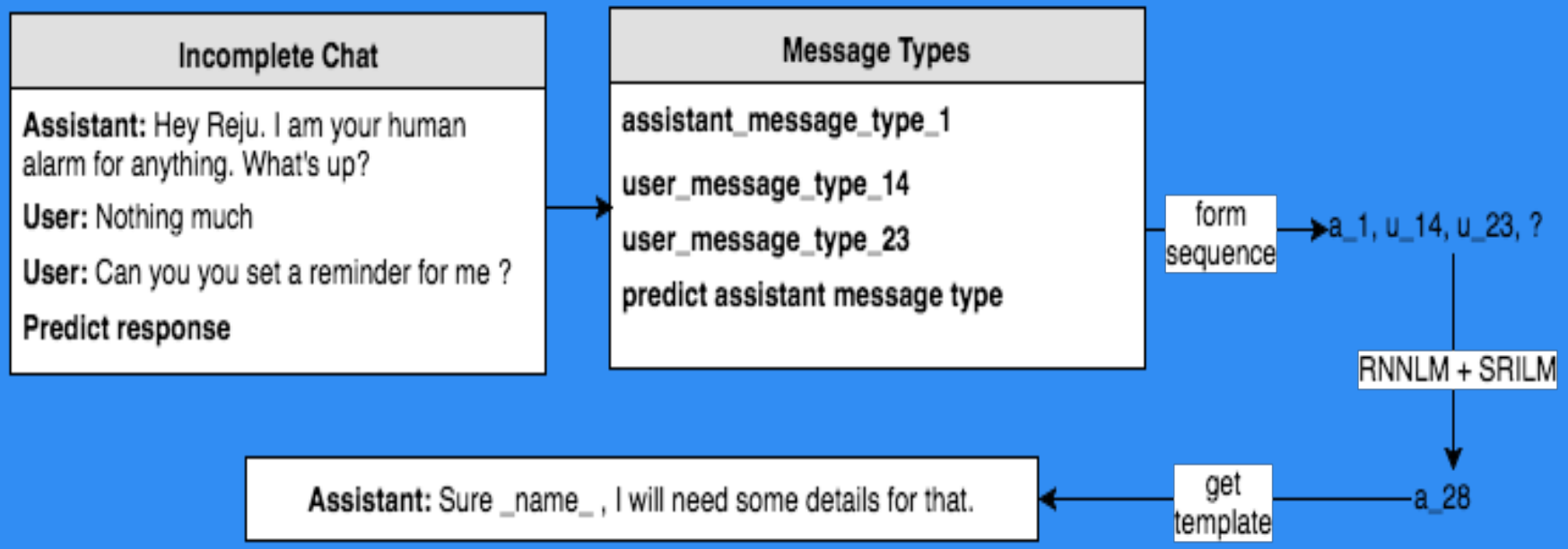
Heuristics, Tf-idf, N-grams
Word Embedding
Word2Vec, Doc2vec,
Document-term Matrix ?



Training Algorithms

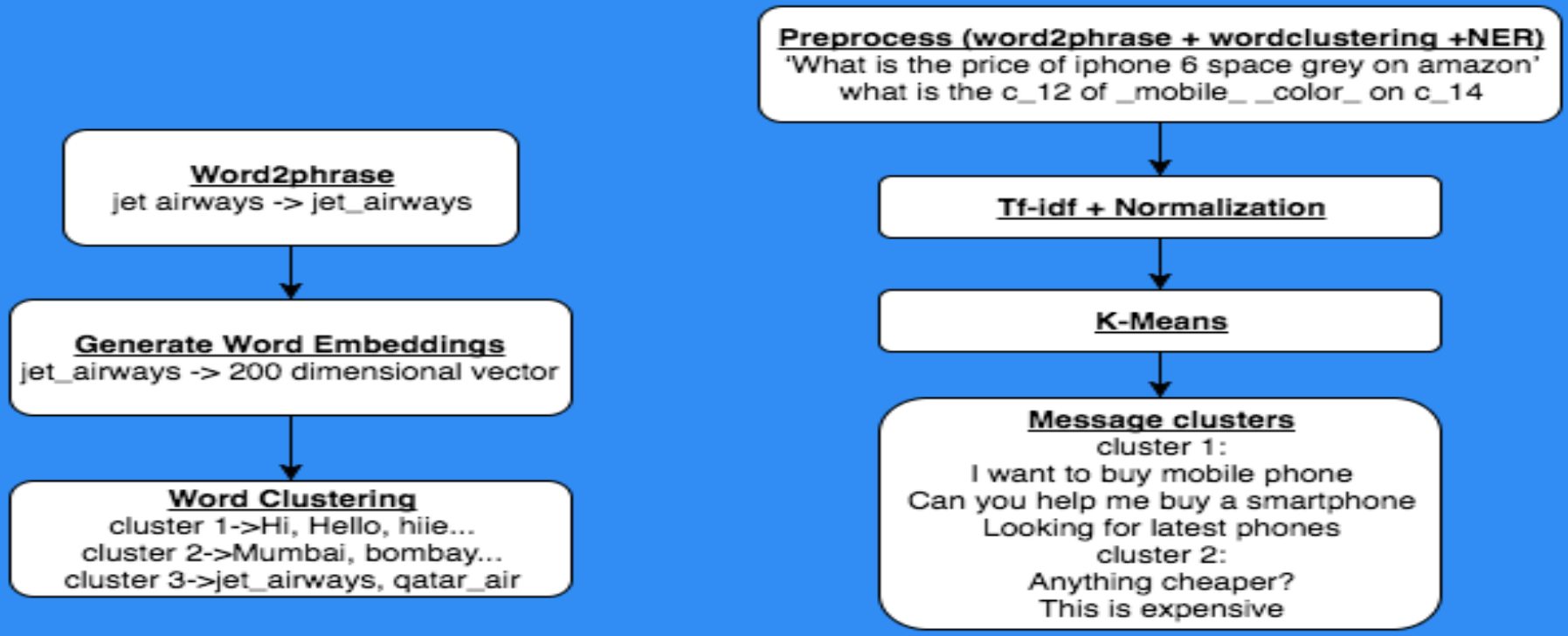
Algorithm	Graph/ Finite State Machine	Retrieval Approach	Generative Approach
Fundamentals	• Graph Traversal (BFS/DFS) • Selects from predefined responses	• Classification (Random Forest, SVM, RNN, Ensemble) • Selects from predefined responses	• Machine Translation (Sequence to Sequence) • Generates its own Natural Language Response
Pros	• Precision is very high • Grammar and personality • Works well for narrow domain seen data	• Precision can be controlled • Grammar and personality • Works well for Narrow domain • Can take complex features	• High Recall • 'SMARTER' – captures context • Caters to unseen Data
Cons	• Fails for unseen data • Capturing context is difficult • Cannot take complex features • Not scalable for open domains	• Capturing context is difficult • Needs lot of iterations during feature extraction • Not Scalable for open domains	• Difficult to define hard precision boundary • Incorrect grammar • Inaccurate for open domains • Needs huge amount of Data

Sequence Learning Example



Overview of Sequence Learning

Clustering of Messages



Message Clustering

Challenges

- Remembering Context – Overwrite/ Replace ?
- Dynamic Information – Date, Time, Prices, Festivals
- Coherent Personality
- Evaluation Metrics
- Diversity of responses

Thank You