(AI for Text Analytics)

問答系統與對話系統
(Question Answering and Dialogue Systems)

1091AITA12
MBA, IMTKU (M2455) (8418) (Fall 2020)
Thu 3, 4 (10:10-12:00) (B206)

Min-Yuh Day
Associate Professor

Institute of Information Management, National Taipei University

https://web.ntpu.edu.tw/~myday
2020-12-31
課程大綱 (Syllabus)

週次 (Week)  日期 (Date)  內容 (Subject/Topics)
1 2020/09/17 人工智慧文本分析課程介紹
   (Course Orientation on Artificial Intelligence for Text Analytics)
2 2020/09/24 文本分析的基礎：自然語言處理
   (Foundations of Text Analytics: Natural Language Processing; NLP)
3 2020/10/01 中秋節 (Mid-Autumn Festival) 放假一天 (Day off)
4 2020/10/08 Python自然語言處理
   (Python for Natural Language Processing)
5 2020/10/15 處理和理解文本
   (Processing and Understanding Text)
6 2020/10/22 文本表達特徵工程
   (Feature Engineering for Text Representation)
課程大綱 (Syllabus)

週次 (Week) 日期 (Date)  內容 (Subject/Topics)
7 2020/10/29  人工智慧文本分析個案研究 I  
              (Case Study on Artificial Intelligence for Text Analytics I)
8 2020/11/05  文本分類  
              (Text Classification)
9 2020/11/12  文本摘要和主題模型  
              (Text Summarization and Topic Models)
10 2020/11/19  期中報告 (Midterm Project Report)
11 2020/11/26  文本相似度和分群  
              (Text Similarity and Clustering)
12 2020/12/03  語意分析和命名實體識別  
              (Semantic Analysis and Named Entity Recognition; NER)
課程大綱 (Syllabus)

週次 (Week)  日期 (Date)  內容 (Subject/Topics)
13 2020/12/10  情感分析  (Sentiment Analysis)
14 2020/12/17  人工智慧文本分析個案研究 II  (Case Study on Artificial Intelligence for Text Analytics II)
15 2020/12/24  深度學習和通用句子嵌入模型  (Deep Learning and Universal Sentence-Embedding Models)
16 2020/12/31  問答系統與對話系統  (Question Answering and Dialogue Systems)
17 2021/01/07  期末報告 I (Final Project Presentation I)
18 2021/01/14  期末報告 II (Final Project Presentation II)
AI for Text Analytics

Text Mining “Knowledge Discovery in Textual Data”

- Document Matching
- Link Analysis
- Search Engines
- Information Retrieval
- POS Tagging
- Lemmatization
- Word Disambiguation

Natural Language Processing

Web Mining
- Web Content Mining
- Web Structure Mining
- Web Usage Mining

Data Mining
- Classification
- Clustering
- Association

Statistics
Machine Learning
Management Science
Artificial Intelligence
Computer Science
Other Disciplines

Question Answering and Dialogue Systems
Outline

• Question Answering

• Dialogue Systems

• Task Oriented Dialogue System
AIWISFIN
AI Conversational Robo-Advisor
(人工智慧對話式理財機器人)
First Place, InnoServe Awards 2018

https://www.youtube.com/watch?v=sEhmyoTXmGk
2018 The 23\textsuperscript{th} International ICT Innovative Services Awards (InnoServe Awards 2018)

- Annual ICT application competition held for university and college students
- The largest and the most significant contest in Taiwan.
- More than ten thousand teachers and students from over one hundred universities and colleges have participated in the Contest.

https://innoserve.tca.org.tw/award.aspx
2018 International ICT Innovative Services Awards (InnoServe Awards 2018)
(2018第23屆大專校院資訊應用服務創新競賽)

https://innoserve.tca.org.tw/award.aspx
IMTKU
Emotional Dialogue System
for
Short Text Conversation
at
NTCIR-14 STC-3 (CECG) Task
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-9 RITE

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Chun Tu

myday@mail.tku.edu.tw

NTCIR-9 Workshop, December 6-9, 2011, Tokyo, Japan
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-10 RITE-2

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Chun Tu
Hou-Cheng Vong
Shih-Wei Wu
Shih-Jhen Huang

myday@mail.tku.edu.tw
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-11 RITE-VAL

Tamkang University

2014

Min-Yuh Day
Ya-Jung Wang
Che-Wei Hsu
En-Chun Tu

Huai-Wen Hsu
Yu-An Lin
Shang-Yu Wu
Yu-Hsuan Tai
Cheng-Chia Tsai

NTCIR-11 Conference, December 8-12, 2014, Tokyo, Japan
IMTKU Question Answering System for World History Exams at NTCIR-12 QA Lab2

Department of Information Management
Tamkang University, Taiwan

myday@mail.tku.edu.tw

NTCIR-12 Conference, June 7-10, 2016, Tokyo, Japan
IMTKU Question Answering System for World History Exams at NTCIR-13 QALab-3

Department of Information Management
Tamkang University, Taiwan

myday@mail.tku.edu.tw

NTCIR-13 Conference, December 5-8, 2017, Tokyo, Japan
IMTKU Emotional Dialogue System for Short Text Conversation at NTCIR-14 STC-3 (CECG) Task

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day  Chi-Sheng Hung  Yi-Jun Xie  Jhiih-Yi Chen  Yu-Ling Kuo  Jian-Ting Lin

myday@mail.tku.edu.tw

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
## 2020 NTCIR-15 Dialogue Evaluation (DialEval-1) Task

### Dialogue Quality (DQ) and Nugget Detection (ND)

#### Chinese Dialogue Quality (S-score) Results (Zeng et al., 2020)

<table>
<thead>
<tr>
<th>Run</th>
<th>Mean RSNOD</th>
<th>Run</th>
<th>Mean NMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMTKU-run2</td>
<td>0.1918</td>
<td>IMTKU-run2</td>
<td>0.1254</td>
</tr>
<tr>
<td>IMTKU-run1</td>
<td>0.1964</td>
<td>IMTKU-run0</td>
<td>0.1284</td>
</tr>
<tr>
<td>IMTKU-run0</td>
<td>0.1977</td>
<td>IMTKU-run1</td>
<td>0.1290</td>
</tr>
<tr>
<td>TUA1-run2</td>
<td>0.2024</td>
<td>TUA1-run2</td>
<td>0.1310</td>
</tr>
<tr>
<td>TUA1-run0</td>
<td>0.2053</td>
<td>TUA1-run0</td>
<td>0.1322</td>
</tr>
<tr>
<td>NKUST-run1</td>
<td>0.2057</td>
<td>NKUST-run1</td>
<td>0.1363</td>
</tr>
<tr>
<td>BL-lstm</td>
<td>0.2088</td>
<td>TUA1-run1</td>
<td>0.1397</td>
</tr>
<tr>
<td>WUST-run0</td>
<td>0.2131</td>
<td>BL-popularity</td>
<td>0.1442</td>
</tr>
<tr>
<td>RSLNV-run0</td>
<td>0.2141</td>
<td>BL-lstm</td>
<td>0.1455</td>
</tr>
<tr>
<td>BL-popularity</td>
<td>0.2288</td>
<td>RSLNV-run0</td>
<td>0.1483</td>
</tr>
<tr>
<td>TUA1-run1</td>
<td>0.2302</td>
<td>WUST-run0</td>
<td>0.1540</td>
</tr>
<tr>
<td>NKUST-run0</td>
<td>0.2653</td>
<td>NKUST-run0</td>
<td>0.2289</td>
</tr>
<tr>
<td>BL-uniform</td>
<td>0.2811</td>
<td>BL-uniform</td>
<td>0.2497</td>
</tr>
</tbody>
</table>

IMTKU System Architecture for NTCIR-13 QALab-3

Question (XML)

Question Analysis

Document Retrieval

Answer Extraction

Answer Generation

Answer (XML)

- Complex Essay
- Simple Essay
- True-or-False
- Factoid
- Slot-Filling
- Unique

- JA&EN Translator
- Stanford CoreNLP
- Wikipedia

- Word Embedding
  - Wiki Word2Vec

NTCIR-13 Conference, December 5-8, 2017, Tokyo, Japan
System Architecture of Intelligent Dialogue and Question Answering System

- User Question Input
- Dialogue Intention Detection
  - RNN LSTM GRU
  - AIML Dialogue Engine
  - Real Time Dialogue API
    - AIML KB
    - Cloud Resource
  - System Response Generator

- Question Analysis
  - Python NLTK
- Document Retrieval
  - Deep Learning TensorFlow
  - Dialogue KB
  - IR
- Answer Extraction
  - Deep Learning
- Answer Generation
- Answer Validation
- Answer

- System Response Generator
IMTKU Emotional Dialogue System Architecture

1. Retrieval-Based Model
2. Generation-Based Model
3. Emotion Classification Model
4. Response Ranking
The system architecture of IMTKU retrieval-based model for NTCIR-14 STC-3

Retrieval-Based Model

1.

Post

Word Segmentation

Keyword Boolean Query

Corpus

Building Index

Retrieval Model

Distinct Result Data

Emotion Matching

Emotion Classification

Word2Vec Similarity Ranking

Retrieval-Based Response

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
The system architecture of IMTKU generation-based model for NTCIR-14 STC-3

**Generation-Based Model**

1. **Training Data**
   - Building Word Index
   - Word Embedding
   - Training Data Seq2seq model

2. **Post**
   - Word Segmentation
   - Short Text Emotion Classifier
   - Trained Model
   - Emotion Matching
   - Word2Vec Similarity Ranking

3. **Generation-Based Response**

**Generative Model**
The system architecture of IMTKU emotion classification model for NTCIR-14 STC-3

Emotion Classification Model

Corpus → Emotion Classification → Training Dataset → MLP, LSTM, BiLSTM → Testing Dataset → Emotion Classification Model → Emotion Prediction
The system architecture of IMTKU Response Ranking for NTCIR-14 STC-3

Response Ranking

1. STC3 Corpus
2. Chinese Segmentation using Jieba
3. Stop Words Removal
4. Word2Vec
5. 1.2 million data (300 dimensions)
6. Vector of Corpus
Short Text Conversation Task (STC-3) 
Chinese Emotional Conversation Generation (CECG) Subtask

Source: http://coai.cs.tsinghua.edu.cn/hml/challenge.html
# NTCIR Short Text Conversation

**STC-1, STC-2, STC-3**

<table>
<thead>
<tr>
<th></th>
<th>Japanese</th>
<th>Chinese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTCIR-12 STC-1</td>
<td>Twitter, Retrieval</td>
<td>Weibo, Retrieval</td>
<td></td>
</tr>
<tr>
<td>22 active participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTCIR-13 STC-2</td>
<td>Yahoo! News, Retrieval+Generation</td>
<td>Weibo, Retrieval+Generation</td>
<td></td>
</tr>
<tr>
<td>27 active participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTCIR-14 STC-3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Chinese Emotional Conversation Generation (CECG) subtask**

**Dialogue Quality (DQ) and Nugget Detection (ND) subtasks**

**Single-turn, Non task-oriented**

**Multi-turn, task-oriented (helpdesk)**

Source: [https://waseda.app.box.com/v/STC3atNTCIR-14](https://waseda.app.box.com/v/STC3atNTCIR-14)
# Chatbots: Evolution of UI/UX

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>mid - 80s</th>
<th>mid - 90s</th>
<th>mid - 00s</th>
<th>mid - 10s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC</td>
<td>Web</td>
<td>Smartphone</td>
<td>Messaging</td>
</tr>
<tr>
<td>Platform</td>
<td>Desktop</td>
<td>Browser</td>
<td>Mobile OS</td>
<td>Messaging Apps</td>
</tr>
<tr>
<td></td>
<td>DOS, Windows, Mac OS</td>
<td>Mosaic, Explorer, Chrome</td>
<td>iOS, Android</td>
<td>WhatsApp, Messenger, Slack</td>
</tr>
<tr>
<td>Applications</td>
<td>Clients</td>
<td>Website</td>
<td>Apps</td>
<td>Bots</td>
</tr>
<tr>
<td></td>
<td>Excel, PPT, Lotus</td>
<td>Yahoo, Amazon</td>
<td>Angry Birds, Instagram</td>
<td>Weather, Travel</td>
</tr>
<tr>
<td>UI/UX</td>
<td>Native Screens</td>
<td>Web Pages</td>
<td>Native Mobile Screens</td>
<td>Message</td>
</tr>
<tr>
<td>S/w Dev</td>
<td>Client-side</td>
<td>Server-side</td>
<td>Client-side</td>
<td>Server-side</td>
</tr>
</tbody>
</table>

Source: https://bbvaopen4u.com/en/actualidad/want-know-how-build-conversational-chatbot-here-are-some-tools
AI Humanoid
Robo-Advisor
AI Humanoid Robo-Advisor for Multi-channel Conversational Commerce

AI Portfolio Asset Allocation

AI Conversation Dialog System

Multichannel Platforms
- Web
- LINE
- Facebook
- Humanoid Robot
System Architecture of AI Humanoid Robo-Advisor

Asset Allocation
- Get daily trading info (Date & AdjClose)
- Data Preprocessing
- Build forecasting model
- Create Portfolio
- Strategy comparison
  - Black-Litterman
  - Markowitz
  - Average Weight
  - Buy and hold

Model Training
- Set up parameters
- Construct model
- Model testing
- Model Comparison

Dialog System
- Sequence to Sequence
  - Use STC _ Weibo for model training
  - Remove words with low frequency
- Data Preprocessing
- Build Seq2Seq model
- Train Seq2Seq
- Evaluate
- Set up the parameters of LSTM and start to train

Dialog System
- AIML
  - Get financial Q&A
  - Convert Q&A information into AIML tag
- Build up our knowledge base

Platforms
- LINE
- FACEBOOK
- Web Application
- ALPHA 2 (AI Humanoid Robo-advisor)
Conversational Model
(LINE, FB Messenger)
Conversational Robo-Advisor
Multichannel UI/UX
Robots

ALPHA 2

ZENBO
AI Dialogue System
Chatbot
Dialogue System
Intelligent Agent
Chatbot

Source: https://www.mdsdecoded.com/blog/the-rise-of-chatbots/
Dialogue System

Overall Architecture of Intelligent Chatbot

Can machines think? (Alan Turing, 1950)

Chatbot

“online human-computer dialog system with natural language.”

Chatbot Conversation Framework

- **Open Domain**
  - Impossible [Hardest]
  - Smart Machine [Hard]

- **Closed Domain**
  - Rules-Based [Easiest]
  - Generative-Based

**Responses**

Source: https://chatbotslife.com/ultimate-guide-to-leveraging-nlp-machine-learning-for-you-chatbot-531ff2dd870c
Chatbots

Bot Maturity Model

Customers want to have simpler means to interact with businesses and get faster response to a question or complaint.

From E-Commerce to Conversational Commerce: Chatbots and Virtual Assistants

Source: http://www.guided-selling.org/from-e-commerce-to-conversational-commerce/
Conversational Commerce: eBay AI Chatbots

![Image of eBay chatbot conversation]

Hotel Chatbot

**BookHotel**

- I’d like to book a hotel
- Sure, which city?
- New York City
- What date are you leaving?
- November 30th, 2016
- Are you sure you want to book the hotel in NYC?
- Yes
- Thank you. The reservation went through successfully.

**Intents**
An intent performs an action in response to natural language user input.

**Utterances**
Spoken or typed phrases that invoke your intent.

**Slots**
Slots are input data required to fulfill the intent.

**Fulfillment**
Fulfillment mechanism for your intent.

Source: https://sdtimes.com/amazon/guest-view-capitalize-amazon-lex-available-general-public/
H&M’s Chatbot on Kik

Source: http://www.guided-selling.org/from-e-commerce-to-conversational-commerce/
Uber’s Chatbot on Facebook’s Messenger

Uber’s chatbot on Facebook’s messenger - one main benefit: it loads much faster than the Uber app

Source: http://www.guided-selling.org/from-e-commerce-to-conversational-commerce/
Savings Bot

Mastercard Makes Commerce More Conversational

Bot Life Cycle and Platform Ecosystem
The Bot Lifecycle

Source: https://chatbotsmagazine.com/the-bot-lifecycle-1ff357430db7
The bot platform ecosystem
and the emerging giants

Nearly every large software company has announced some sort of bot strategy in the last year. Here’s a look at a handful of leading platforms that developers might use to send messages, interpret natural language, and deploy bots, with the emerging bot-ecosystem giants highlighted.

General AI agents with platforms
Developer access available now or announced

Source: https://www.oreilly.com/ideas/infographic-the-bot-platform-ecosystem
Bot frameworks and deployment platforms

Wit.ai
Facebook

BotKit
Howdy

Chatfuel

Automat

Bot Framework
Microsoft

Api.ai
Google

Pandorabots

MindMeld

Gupshup

Sequel

Source: https://www.oreilly.com/ideas/infographic-the-bot-platform-ecosystem
Messenger Bot Landscape

May 2017

Food
- The Wise Parent
- Plum
- Pretzelmaker Kitchen
- Hungry
- Foodie

Communication
- TangoWork
- Typeform
- Away
- Trujillo
- Refugg
- Rescue
- Messenger Match

Utilities
- Pancho
- CarBot
- Smoke
- DotCom
- Serve Messenger

Personal
- M
- Assif
- Operator
- Uber
- Swiftly
- AskAnia

Design
- ColorBot
- Connie Digital
- AWWWARDS
- Mr. Norman
- Graphic Design
- SnailBot

Analytics
- SENSE
- Sizzle
- PopDaily
- DMT
- BuzzLogger
- Trading Bot

News
- CNN
- RT
- Digg
- WSJ
- Reddit Bot
- Al-Jazeera

Travel
- Gobos
- KLM
- British Airways
- Spark Ego
- Austrian Airlines

Entertainment
- Spotify
- Kim Kardashian
- La Bragade
- 50 Cent
- CocaCola
- Lindsey Lohan
- Marcus S

Developer Tools
- HackerOne
- Winedia
- Rebie
- Zry

Education
- Genius
- Kichchi

Source: https://medium.com/@RecastAI/2017-messenger-bot-landscape-a-public-spreadsheet-gathering-1000-messenger-bots-f017fdb1448a
How to Build Chatbots

END USER
sends/receives message

MESSAGING PLATFORMS
requests webhook URL / receives response

WEBHOOK

DATABASE
query data from DB

APIs
interact with API for data

BOT APPLICATION
server handles requests via CGI

WEB SERVER
APACHE
NGINX

Chatbot Frameworks and AI Services

• Bot Frameworks
  – Botkit
  – Microsoft Bot Framework
  – Rasa NLU

• AI Services
  – Wit.ai
  – api.ai
  – LUIS.ai
  – IBM Watson

Chatbot Frameworks

Comparison Table of Most Prominent Bot Frameworks

<table>
<thead>
<tr>
<th>Feature</th>
<th>Botkit</th>
<th>Microsoft Bot Framework</th>
<th>RASA NLU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Built-In integration with messaging platforms</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td>NLP support</td>
<td>❌</td>
<td>❌ but have close bonds with LUIS.ai</td>
<td>✔️</td>
</tr>
<tr>
<td>Out-of-box bots ready to be deployed</td>
<td>✔️</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td>Programming Language</td>
<td>JavaScript (Node)</td>
<td>JavaScript (Node), C#</td>
<td>Python</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison of Most Prominent AI Services</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>wit.ai</strong></td>
</tr>
<tr>
<td>Free of charge</td>
</tr>
<tr>
<td>Text and Speech processing</td>
</tr>
<tr>
<td>Machine Learning Modeling</td>
</tr>
<tr>
<td>Support for Intents, Entities, Actions</td>
</tr>
<tr>
<td>Support entities, actions are combined operations</td>
</tr>
<tr>
<td>Pre-build entities for easy parsing of numbers, temperature, date, etc.</td>
</tr>
<tr>
<td>Integration to messaging platforms</td>
</tr>
<tr>
<td>Support of SDKs</td>
</tr>
<tr>
<td>includes SDKs for Python, Node.js, Rust, C, Ruby, iOS, Android, Windows Phone</td>
</tr>
</tbody>
</table>

Transformer (Attention is All You Need) (Vaswani et al., 2017)

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT (Bidirectional Encoder Representations from Transformers)

Overall pre-training and fine-tuning procedures for BERT

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT (Bidirectional Encoder Representations from Transformers)

BERT input representation

BERT, OpenAI GPT, ELMo

Fine-tuning BERT on Different Tasks

(a) Sentence Pair Classification Tasks:
- MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

(b) Single Sentence Classification Tasks:
- SST-2, CoLA

(c) Question Answering Tasks:
- SQuAD v1.1

(d) Single Sentence Tagging Tasks:
- CoNLL-2003 NER

Fine-tuning BERT on Question Answering (QA)

(c) Question Answering Tasks: SQuAD v1.1

Fine-tuning BERT on Dialogue Intent Detection (ID; Classification)

(b) Single Sentence Classification Tasks: SST-2, CoLA

Fine-tuning BERT on Dialogue Slot Filling (SF)

Pre-trained Language Model (PLM)

Source: https://github.com/thunlp/PLMpapers
Turing Natural Language Generation (T-NLG)

Transformers

State-of-the-art Natural Language Processing for TensorFlow 2.0 and PyTorch

- Transformers
  - pytorch-transformers
  - pytorch-pretrained-bert
- provides state-of-the-art general-purpose architectures
  - (BERT, GPT-2, RoBERTa, XLM, DistilBert, XLNet, CTRL...)
  - for Natural Language Understanding (NLU) and Natural Language Generation (NLG)
  - with over 32+ pretrained models in 100+ languages and deep interoperability between TensorFlow 2.0 and PyTorch.

Source: https://github.com/huggingface/transformers
Transfer Learning in Natural Language Processing

# NLP Benchmark Datasets

<table>
<thead>
<tr>
<th>Task</th>
<th>Dataset</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WMT 2014 EN-FR</td>
<td></td>
</tr>
<tr>
<td>Text Summarization</td>
<td>CNN/DM</td>
<td><a href="https://cs.nyu.edu/~kcho/DMQA/">https://cs.nyu.edu/~kcho/DMQA/</a></td>
</tr>
<tr>
<td></td>
<td>Newsroom</td>
<td><a href="https://summari.es/">https://summari.es/</a></td>
</tr>
<tr>
<td></td>
<td>Gigaword</td>
<td><a href="https://catalog.ldc.upenn.edu/LDC2012T21">https://catalog.ldc.upenn.edu/LDC2012T21</a></td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>ARC</td>
<td><a href="http://data.allenai.org/arc/">http://data.allenai.org/arc/</a></td>
</tr>
<tr>
<td>Question Generation</td>
<td>CNN/DM</td>
<td><a href="https://cs.nyu.edu/~kcho/DMQA/">https://cs.nyu.edu/~kcho/DMQA/</a></td>
</tr>
<tr>
<td></td>
<td>NewsQA</td>
<td><a href="https://datasets.maluuba.com/NewsQA">https://datasets.maluuba.com/NewsQA</a></td>
</tr>
<tr>
<td></td>
<td>SQuAD</td>
<td><a href="https://rajpurkar.github.io/SQuAD-explorer/">https://rajpurkar.github.io/SQuAD-explorer/</a></td>
</tr>
<tr>
<td></td>
<td>NarrativeQA</td>
<td><a href="https://github.com/deepmind/narrativeqa">https://github.com/deepmind/narrativeqa</a></td>
</tr>
<tr>
<td></td>
<td>Quasar</td>
<td><a href="https://github.com/bdhingra/quasar">https://github.com/bdhingra/quasar</a></td>
</tr>
<tr>
<td></td>
<td>SearchQA</td>
<td><a href="https://github.com/nyu-dl/SearchQA">https://github.com/nyu-dl/SearchQA</a></td>
</tr>
<tr>
<td>Semantic Parsing</td>
<td>AMR parsing</td>
<td><a href="https://amr.isi.edu/index.html">https://amr.isi.edu/index.html</a></td>
</tr>
<tr>
<td></td>
<td>ATIS (SQL Parsing)</td>
<td><a href="https://github.com/jkkummerfeld/text2sql-data/tree/master/data">https://github.com/jkkummerfeld/text2sql-data/tree/master/data</a></td>
</tr>
<tr>
<td></td>
<td>WikiSQL (SQL Parsing)</td>
<td><a href="https://github.com-salesforce/wikiSQL">https://github.com-salesforce/wikiSQL</a></td>
</tr>
<tr>
<td></td>
<td>SST</td>
<td><a href="https://nlp.stanford.edu/sentiment/index.html">https://nlp.stanford.edu/sentiment/index.html</a></td>
</tr>
<tr>
<td></td>
<td>Yelp Reviews</td>
<td><a href="https://www.yelp.com/dataset/challenge">https://www.yelp.com/dataset/challenge</a></td>
</tr>
<tr>
<td></td>
<td>DBpedia</td>
<td><a href="https://wiki.dbpedia.org/Datasets">https://wiki.dbpedia.org/Datasets</a></td>
</tr>
<tr>
<td></td>
<td>TREC</td>
<td><a href="https://trec.nist.gov/data.html">https://trec.nist.gov/data.html</a></td>
</tr>
<tr>
<td>Natural Language Inference</td>
<td>SNLI Corpus</td>
<td><a href="https://nlp.stanford.edu/projects/snli/">https://nlp.stanford.edu/projects/snli/</a></td>
</tr>
<tr>
<td></td>
<td>MultiNLI</td>
<td><a href="https://www.nyu.edu/projects/bowman/multiNLI/">https://www.nyu.edu/projects/bowman/multiNLI/</a></td>
</tr>
<tr>
<td></td>
<td>OneNotes</td>
<td><a href="https://catalog.ldc.upenn.edu/LDC2013T19">https://catalog.ldc.upenn.edu/LDC2013T19</a></td>
</tr>
</tbody>
</table>

Question Answering (QA) SQuAD

Stanford Question Answering Dataset
What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or span, from the corresponding reading passage, or the question might be unanswerable.

SQuAD2.0 combines the 100,000 questions in SQuAD1.1 with over 50,000 unanswerable questions written adversarially by crowdworkers to look similar to answerable ones. To do well on SQuAD2.0, systems must not only answer questions when possible, but also determine when no answer is supported by the paragraph and abstain from answering.

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Model</th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Human Performance Stanford University (Rajpurkar &amp; Jia et al. '18)</td>
<td>86.831</td>
<td>89.452</td>
</tr>
<tr>
<td>1</td>
<td>SA-Net on Albert (ensemble) QIANXIN</td>
<td>90.724</td>
<td>93.011</td>
</tr>
<tr>
<td>2</td>
<td>SA-Net-V2 (ensemble) QIANXIN</td>
<td>90.679</td>
<td>92.948</td>
</tr>
<tr>
<td>2</td>
<td>Retro-Reader (ensemble)</td>
<td>90.578</td>
<td>92.978</td>
</tr>
</tbody>
</table>

https://rajpurkar.github.io/SQuAD-explorer/
Abstract

We present the Stanford Question Answering Dataset (SQuAD), a new reading comprehension dataset consisting of 100,000+ questions posed by crowdworkers on a set of Wikipedia articles, where the answer to each question is a segment of text from the corresponding reading passage. We analyze the dataset to understand the types of reasoning required to answer the questions, leaning heavily on dependency and constituency trees. We build a strong logistic regression model, which achieves an F1 score of 51.0%, a significant improvement over a simple baseline (20%). However, human performance (86.8%) is much higher, indicating that the dataset presents a good challenge problem for future research. The dataset is freely available at https://stanford-qa.com.

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

What causes precipitation to fall?
gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
graupel

Where do water droplets collide with ice crystals to form precipitation?
within a cloud

Figure 1: Question-answer pairs for a sample passage in the
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity from clouds. The main forms of precipitation include drizzle, rain, sleet, snow, ice pellets, graupel and hail. Precipitation occurs when a portion of the atmosphere becomes saturated with water vapor (reaching 100% relative humidity), so that the water condenses and "precipitates". Thus, fog and mist are not precipitation but suspensions, because the water vapor does not condense sufficiently to precipitate. Two processes, possibly acting together, can lead to air becoming saturated: cooling the air or adding water vapor to the air. Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called "showers".
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

**Q:** What causes precipitation to fall?

**A:** gravity
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
A: graupel
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called “showers”.

**Q:** Where do water droplets collide with ice crystals to form precipitation?

**A:** **within a cloud**
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?
A: gravity

Q: What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
A: graupel

Q: Where do water droplets collide with ice crystals to form precipitation?
A: within a cloud
Super Bowl 50

From Wikipedia, the free encyclopedia

"2016 Super Bowl" redirects here. For the Super Bowl that was played at the completion of the 2016 season, see Super Bowl LI.

"SB 50" redirects here. For the California transit-density bill, see California Senate Bill 50.

Super Bowl 50 was an American football game to determine the champion of the National Football League (NFL) for the 2015 season. The American Football Conference (AFC) champion Denver Broncos defeated the National Football Conference (NFC) champion Carolina Panthers, 24–10. The game was played on February 7, 2016, at Levi's Stadium in Santa Clara, California, in the San Francisco Bay Area. As this was the 50th Super Bowl game, the league emphasized the "golden anniversary" with various gold-themed initiatives during the 2015 season, as well as suspending the tradition of naming each Super Bowl game with Roman numerals (under which the game would have been known as "Super Bowl L"), so the logo could prominently feature the Arabic numerals 5 and 0.\[^5\][^6]

The Panthers finished the regular season with a 15–1 record, racking up the league's top offense, and quarterback Cam Newton was named the NFL Most Valuable Player (MVP). They defeated the Arizona Cardinals 49–15 in the NFC Championship Game and advanced to their second Super Bowl appearance since the franchise began playing in 1995. The Broncos finished the regular season with a 12–4 record, bolstered by having the league's top defense. The Broncos defeated the defending Super Bowl champion New England Patriots 20–18 in the AFC Championship Game joining the Patriots, Dallas Cowboys, and Pittsburgh Steelers as one of four teams that have made eight appearances in the Super Bowl. This record would later be broken the next season, in 2017, when the Patriots advanced to their ninth Super Bowl appearance in Super Bowl LI.

https://en.wikipedia.org/wiki/Super_Bowl_50
Dialogue on Airline Travel Information System (ATIS)
The ATIS (Airline Travel Information System) Dataset


Training samples: 4978
Testing samples: 893
Vocab size: 943
Slot count: 129
Intent count: 26

https://www.kaggle.com/siddhadev/atis-dataset-from-ms-cntk
SF-ID Network (E et al., 2019)
Slot Filling (SF)
Intent Detection (ID)

A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling

Intent Detection on ATIS

State-of-the-art

<table>
<thead>
<tr>
<th>RANK</th>
<th>METHOD</th>
<th>ACCURACY</th>
<th>PAPER TITLE</th>
<th>YEAR</th>
<th>PAPER</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SF-ID</td>
<td>0.9776</td>
<td>A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling</td>
<td>2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Capsule-NLU</td>
<td>0.950</td>
<td>Joint Slot Filling and Intent Detection via Capsule Neural Networks</td>
<td>2018</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: https://paperswithcode.com/sota/intent-detection-on-atis
Slot Filling on ATIS
State-of-the-art

<table>
<thead>
<tr>
<th>RANK</th>
<th>METHOD</th>
<th>F1</th>
<th>PAPER TITLE</th>
<th>YEAR</th>
<th>PAPER</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SF-ID</td>
<td>0.958</td>
<td>A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling</td>
<td>2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Capsule-NLU</td>
<td>0.952</td>
<td>Joint Slot Filling and Intent Detection via Capsule Neural Networks</td>
<td>2018</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: [https://paperswithcode.com/sota/slot-filling-on-atis](https://paperswithcode.com/sota/slot-filling-on-atis)
Restaurants Dialogue Datasets

• MIT Restaurant Corpus
  – https://groups.csail.mit.edu/sls/downloads/restaurant/

• CamRest676
  (Cambridge restaurant dialogue domain dataset)
  – https://www.repository.cam.ac.uk/handle/1810/260970

• DSTC2 (Dialog State Tracking Challenge 2 & 3)
  – http://camdial.org/~mh521/dstc/
CrossWOZ:
A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset

<table>
<thead>
<tr>
<th>MultiWOZ</th>
<th>CrossWOZ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>usr:</strong> I’m looking for a college type attraction.</td>
<td><strong>usr:</strong> 你好，可以帮我推荐一个评分是4.5分以上的景点吗？</td>
</tr>
<tr>
<td>...</td>
<td><strong>sys:</strong> Hello, could you recommend an attraction with a rating of 4.5 or higher?</td>
</tr>
<tr>
<td><strong>usr:</strong> I would like to visit in town centre please.</td>
<td><strong>sys:</strong> 天安门城楼，簋街小吃和北京欢乐谷都是很不错的地方呢。</td>
</tr>
<tr>
<td>...</td>
<td><strong>usr:</strong> 天安门城楼，簋街小吃和北京欢乐谷都是很不错的地方呢。</td>
</tr>
<tr>
<td><strong>usr:</strong> Can you find an Indian restaurant for me that is <em>also in the town centre</em>?</td>
<td><strong>usr:</strong> 我喜欢北京欢乐谷，你知道这个景点周边的酒店都是什么吗？</td>
</tr>
<tr>
<td><strong>Schema</strong></td>
<td><strong>sys:</strong> I like Beijing Happy Valley. What hotels are around this attraction?</td>
</tr>
<tr>
<td><strong>usr:</strong> I want a hotel in <em>San Diego</em> and I want to check out on Thursday next week.</td>
<td><strong>sys:</strong> 那可多了，有A酒店，B酒店，C酒店。</td>
</tr>
<tr>
<td>...</td>
<td><strong>usr:</strong> 那可多了，有A酒店，B酒店，C酒店。</td>
</tr>
<tr>
<td><strong>usr:</strong> I need a one way flight to <em>go there</em>.</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>CrossWOZ</strong></td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>usr:</strong> I’m looking for a college type attraction.</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>usr:</strong> I would like to visit in town centre please.</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>usr:</strong> Can you find an Indian restaurant for me that is <em>also in the town centre</em>?</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>Schema</strong></td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>usr:</strong> I want a hotel in <em>San Diego</em> and I want to check out on Thursday next week.</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td>...</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>usr:</strong> I need a one way flight to <em>go there</em>.</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td><strong>CrossWOZ</strong></td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
</tbody>
</table>

## CrossWOZ: A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset

<table>
<thead>
<tr>
<th>Type</th>
<th>Single-domain goal</th>
<th>Multi-domain goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset</td>
<td>DSTC2</td>
<td>WOZ 2.0</td>
</tr>
<tr>
<td>Language</td>
<td>EN</td>
<td>EN</td>
</tr>
<tr>
<td>Speakers</td>
<td>H2M</td>
<td>H2H</td>
</tr>
<tr>
<td># Domains</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td># Dialogues</td>
<td>1,612</td>
<td>600</td>
</tr>
<tr>
<td># Turns</td>
<td>23,354</td>
<td>4,472</td>
</tr>
<tr>
<td>Avg. domains</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Avg. turns</td>
<td>14.5</td>
<td>7.5</td>
</tr>
<tr>
<td># Slots</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td># Values</td>
<td>212</td>
<td>99</td>
</tr>
</tbody>
</table>

Task-Oriented Dialogue

Initial user state (=user goal)

id=1 (Attraction): fee=free, name=?, nearby hotels=?

id=2 (Hotel): name=near (id=1), wake-up call=yes, rating=?

id=3 (Taxi): from=(id=1), to=(id=2), car type=?, plate number=?

Final user state

id=1 (Attraction): name=Tiananmen Square, fee=free, nearby hotels=[Beijing Capital Hotel, Guidu Hotel Beijing]

id=2 (Hotel): name=Beijing Capital Hotel, wake-up call=yes, rating=4.6

id=3 (Taxi): from=Tiananmen Square, to=Beijing Capital Hotel, car type=#CX, plate number=#CP

- 自然語言理解
  Natural Language Understanding (NLU)
- 對話管理
  Dialog Management (DM)
- 自然語言生成
  Natural Language Generation (NLG)

Question Answering and Dialogue Systems

Question Answering (QA)

BERT for Question Answering

Source: Apoorv Nandan (2020), BERT (from HuggingFace Transformers) for Text Extraction, https://keras.io/examples/nlp/text_extraction_with_bert/

Description: Fine tune pretrained BERT from HuggingFace Transformers on SQuAD.

Introduction

This demonstration uses SQuAD (Stanford Question-Answering Dataset). In SQuAD, an input consists of a question, and a paragraph for context. The goal is to find the span of text in the paragraph that answers the question. We evaluate our performance on this data with the "Exact Match" metric, which measures the percentage of predictions that exactly match any one of the ground-truth answers.

We fine-tune a BERT model to perform this task as follows:

1. Feed the context and the question as inputs to BERT.
2. Take two vectors S and T with dimensions equal to that of hidden states in BERT.
3. Compute the probability of each token being the start and end of the answer span. The probability of a token being the start of the answer is given by a dot product between S and the representatio of the token in the last layer of BERT, followed by a softmax over all tokens. The probability of a token being the end of the answer is compute similarly with the vector T.
4. Fine-tune BERT and learn S and T along the way.

References:

- BERT
- SQuAD

https://tinyurl.com/aintpupython101
Python in Google Colab (Python101)

https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT

Table of contents

RandomizedSearchCV

Sentiment Analysis

Sentiment Analysis - Unsupervised
Lexical

Sentiment Analysis - Supervised
Machine Learning

Sentiment Analysis - Supervised
Deep Learning Models

Sentiment Analysis - Advanced Deep
Learning

Deep Learning and Universal Sentence-Embedding Models

Universal Sentence Encoder (USE)

Universal Sentence Encoder
Multilingual (USEM)

Question Answering and Dialogue Systems

Question Answering (QA)

BERT for Question Answering

Dialogue Systems

Joint Intent Classification and
Slot Filling with Transformers

Data Visualization

Download: 100%

Output Shape

Param #

Connected to

Layer (type)

input_1 (InputLayer)

input_3 (InputLayer)

input_2 (InputLayer)

tf_bert_model (TFBertModel)

start_logit (Dense)

end_logit (Dense)

flatten (Flatten)

flatten_1 (Flatten)

activation_7 (Activation)

activation_8 (Activation)

Total params: 109,483,776
Trainable params: 109,483,776
Non-trainable params: 0

CPU times: user 20.8 s, sys: 7.75 s, total: 28.5 s
Wall time: 1min 42s

https://tinyurl.com/aintpupython101
Python in Google Colab (Python101)

https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT

Dialogue Systems

Joint Intent Classification and Slot Filling with Transformers

The goal of this notebook is to fine-tune a pretrained transformer-based neural network model to convert a user query expressed in English into a representation that is structured enough to be processed by an automated service.

Here is an example of interpretation computed by such a Natural Language Understanding system:

```python
>>> nlu("Book a table for two at Le Ritz for Friday night",
         tokenizer, joint_model, intent_names, slot_names)
```

```python
{
    'intent': 'BookRestaurant',
    'slots': {
        'party_size_number': 'two',
        'restaurant_name': 'Le Ritz',
        'timeRange': 'Friday night'
    }
}
```

Intent classification is a simple sequence classification problem. The trick is to treat the structured knowledge extraction part ("Slot Filling") as token-level classification problem using BIO-annotations:

https://tinyurl.com/aintpupython101
Python in Google Colab (Python101)

https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT

```
1  def show_predictions(text, tokenizer, model, intent_names, slot_names):
2      inputs = tf.constant(tokenizer.encode(text))[None, : ] # batch_size = 1
3      outputs = model(inputs)
4      slot_logits, intent_logits = outputs
5      slot_ids = slot_logits.numpy().argmax(axis=-1)[0, 1:-1]
6      intent_id = intent_logits.numpy().argmax(axis=-1)[0]
7      print("Text:", text)
8      print("Intent:", intent_names[intent_id])
9      print("Slots:")
10     for token, slot_id in zip(tokenizer.tokenize(text), slot_ids):
11        print(f"{token:>10} : {slot_names[slot_id]}")
12     show_predictions("Book a table for two at Le Ritz for Friday night!",
13             tokenizer, joint_model, intent_names, slot_names)
```

Text: Book a table for two at Le Ritz for Friday night!
Intent: BookRestaurant
Slots:
    Book : 0
    a : 0
    table : 0
    for : 0
    two : B-party_size_number
    at : 0
    Le : B-restaurant_name
    R : I-restaurant_name
    ##itz : I-restaurant_name
    for : 0
    Friday : B-timeRange
    night : 0
    : 0

https://tinyurl.com/aintpuppython101
Python in Google Colab (Python101)

https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT

```
# Naive NLU: handling: treat B- and I- the same...
new_slot_name = current_word_slot_name[2:]
if active_slot_name is None:
    active_slot_words.append(word)
    active_slot_name = new_slot_name
elif new_slot_name == active_slot_name:
    active_slot_words.append(word)
else:
    collected_slots[active_slot_name] = ".".join(active_slot_words)
    active_slot_words = [word]
    active_slot_name = new_slot_name
if active_slot_name:
    collected_slots[active_slot_name] = ".".join(active_slot_words)
info["slots"] = collected_slots
return info

def nlu(text, tokenizer, model, intent_names, slot_names):
    inputs = tf.constant(tokenizer.encode(text))[None, :]
    # batch_size = 1
    outputs = model(inputs)
    slot_logits, intent_logits = outputs
    slot_ids = slot_logits.numpy().argmax(axis=-1)[0, 1:-1]
    intent_id = intent_logits.numpy().argmax(axis=-1)[0]
    return decode_predictions(text, tokenizer, intent_names, slot_names,
                               intent_id, slot_ids)
	nlu("Book a table for two at Le Ritz for Friday night",
          tokenizer, joint_model, intent_names, slot_names)
```
## NLP Benchmark Datasets

<table>
<thead>
<tr>
<th>Task</th>
<th>Dataset</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine Translation</td>
<td>WMT 2014 EN-FR</td>
<td></td>
</tr>
<tr>
<td>Text Summarization</td>
<td>CNN/DM</td>
<td><a href="https://cs.nyu.edu/~kcho/DMQA/">https://cs.nyu.edu/~kcho/DMQA/</a></td>
</tr>
<tr>
<td>Text Summarization</td>
<td>Newsroom</td>
<td><a href="https://summari.es/">https://summari.es/</a></td>
</tr>
<tr>
<td>Text Summarization</td>
<td>Gigaword</td>
<td><a href="https://catalog.ldc.upenn.edu/LDC2012T21">https://catalog.ldc.upenn.edu/LDC2012T21</a></td>
</tr>
<tr>
<td>Reading Comprehension Question Answering</td>
<td>ARC</td>
<td><a href="http://data.allenai.org/arc/">http://data.allenai.org/arc/</a></td>
</tr>
<tr>
<td>Reading Comprehension Question Answering</td>
<td>CliCR</td>
<td><a href="http://aclweb.org/anthology/N18-1140">http://aclweb.org/anthology/N18-1140</a></td>
</tr>
<tr>
<td>Reading Comprehension Question Answering</td>
<td>CNN/DM</td>
<td><a href="https://cs.nyu.edu/~kcho/DMQA/">https://cs.nyu.edu/~kcho/DMQA/</a></td>
</tr>
<tr>
<td>Reading Comprehension Question Answering</td>
<td>NewsQA</td>
<td><a href="https://datasets.maluuba.com/NewsQA">https://datasets.maluuba.com/NewsQA</a></td>
</tr>
<tr>
<td>Reading Comprehension Question Answering</td>
<td>RACE</td>
<td><a href="http://www.qizhexie.com/data/RACE_leaderboard">http://www.qizhexie.com/data/RACE_leaderboard</a></td>
</tr>
<tr>
<td>Reading Comprehension Question Answering</td>
<td>SQuAD</td>
<td><a href="https://rajpurkar.github.io/SQuAD_leaderboard">https://rajpurkar.github.io/SQuAD_leaderboard</a></td>
</tr>
<tr>
<td>Question Answering</td>
<td>NarrativeQA</td>
<td><a href="https://github.com/deepmind/narrativeqa">https://github.com/deepmind/narrativeqa</a></td>
</tr>
<tr>
<td>Question Answering</td>
<td>Quasar</td>
<td><a href="https://github.com/nyu-dl/Quasar">https://github.com/nyu-dl/Quasar</a></td>
</tr>
<tr>
<td>Question Answering</td>
<td>SearchQA</td>
<td></td>
</tr>
<tr>
<td>Semantic Parsing</td>
<td>AMR parsing</td>
<td><a href="http://ai.stanford.edu/index.html">http://ai.stanford.edu/index.html</a></td>
</tr>
<tr>
<td>Semantic Parsing</td>
<td>WikiSQL (SQL Parsing)</td>
<td><a href="https://github.com/salesforce/WikiSQL">https://github.com/salesforce/WikiSQL</a></td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>IMDB Reviews</td>
<td><a href="http://nlp.stanford.edu/sentiment/index.html">http://nlp.stanford.edu/sentiment/index.html</a></td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>SST</td>
<td><a href="https://www.yelp.com/dataset/challenge">https://www.yelp.com/dataset/challenge</a></td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>Yelp Reviews</td>
<td><a href="http://www.cs.cornell.edu/people/pabo/movie-review-data/">http://www.cs.cornell.edu/people/pabo/movie-review-data/</a></td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>Subjectivity Dataset</td>
<td></td>
</tr>
<tr>
<td>Text Classification</td>
<td>DBpedia</td>
<td><a href="https://wiki.dbpedia.org/Datasets">https://wiki.dbpedia.org/Datasets</a></td>
</tr>
<tr>
<td>Text Classification</td>
<td>TREC</td>
<td><a href="https://trec.nist.gov/data.html">https://trec.nist.gov/data.html</a></td>
</tr>
<tr>
<td>Natural Language Inference</td>
<td>SNLI Corpus</td>
<td><a href="https://nlp.stanford.edu/projects/snli/">https://nlp.stanford.edu/projects/snli/</a></td>
</tr>
<tr>
<td>Natural Language Inference</td>
<td>MultiNLI</td>
<td><a href="https://www.nyu.edu/projects/bowman/multiNLI/">https://www.nyu.edu/projects/bowman/multiNLI/</a></td>
</tr>
<tr>
<td>Natural Language Inference</td>
<td>SciTail</td>
<td><a href="http://data.allenai.org/scitail/">http://data.allenai.org/scitail/</a></td>
</tr>
<tr>
<td>Semantic Role Labeling</td>
<td>Proposition Bank</td>
<td><a href="http://propbank.github.io/">http://propbank.github.io/</a></td>
</tr>
<tr>
<td>Semantic Role Labeling</td>
<td>OneNotes</td>
<td><a href="https://catalog.ldc.upenn.edu/LDC2013T19">https://catalog.ldc.upenn.edu/LDC2013T19</a></td>
</tr>
</tbody>
</table>

Summary

• Question Answering

• Dialogue Systems

• Task Oriented Dialogue System
References

- Apoorv Nandan (2020), BERT (from HuggingFace Transformers) for Text Extraction, https://keras.io/examples/nlp/text_extraction_with_bert/
- Olivier Grisel (2020), Transformers (BERT fine-tuning): Joint Intent Classification and Slot Filling, https://m2dsupsdlclass.github.io/lectures-labs/