文字探勘 (Text Mining)
問答系統與對話系統 (Question Answering and Dialogue Systems)

1082TM12
MBA, BDABI, TKU (E3611) (8480) (Spring 2020)
Mon, 7, 8, 9 (14:10-17:00) (B206)

Chichang Jou
周清江
Associate Professor
cjou@mail.tku.edu.tw

Min-Yuh Day
戴敏育
Associate Professor
myday@mail.tku.edu.tw

Dept. of Information Management, Tamkang University
淡江大學 資訊管理學系

2020-06-08
<table>
<thead>
<tr>
<th>週次 (Week)</th>
<th>日期 (Date)</th>
<th>內容 (Subject/Topics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2020/03/02</td>
<td>文字探勘課程介紹  (Course Orientation on Text Mining)</td>
<td></td>
</tr>
<tr>
<td>2 2020/03/09</td>
<td>文字探勘基礎：自然語言處理  (Foundations of Text Mining: Natural Language Processing; NLP)</td>
<td></td>
</tr>
<tr>
<td>3 2020/03/16</td>
<td>Python自然語言處理  (Python for Natural Language Processing)</td>
<td></td>
</tr>
<tr>
<td>4 2020/03/23</td>
<td>處理和理解文本 (Processing and Understanding Text)</td>
<td></td>
</tr>
<tr>
<td>5 2020/03/30</td>
<td>文本表達特徵工程 (Feature Engineering for Text Representation)</td>
<td></td>
</tr>
<tr>
<td>6 2020/04/06</td>
<td>人工智慧文本分析個案研究 I (Case Study on Artificial Intelligence for Text Analytics I)</td>
<td></td>
</tr>
</tbody>
</table>
課程大綱 (Syllabus)

週次 (Week) 日期 (Date)  內容 (Subject/Topics)
7 2020/04/13  2020/04/13 文本分類
               (Text Classification)
8 2020/04/20  2020/04/20 文本摘要和主題模型
               (Text Summarization and Topic Models)
9 2020/04/27  2020/04/27 期中報告 (Midterm Project Report)
10 2020/05/04 2020/05/04 文本相似度和分群
                (Text Similarity and Clustering)
11 2020/05/11 2020/05/11 語意分析和命名實體識別
                (Semantic Analysis and Named Entity Recognition; NER)
12 2020/05/18 2020/05/18 情感分析
                (Sentiment Analysis)
課程大綱 (Syllabus)

週次 (Week) 日期 (Date)  內容 (Subject/Topics)

13  2020/05/25  人工智慧文本分析個案研究 II
    (Case Study on Artificial Intelligence for Text Analytics II)

14  2020/06/01  深度學習和通用句子嵌入模型
    (Deep Learning and Universal Sentence-Embedding Models)

15  2020/06/08  問答系統與對話系統
    (Question Answering and Dialogue Systems)

16  2020/06/15  期末報告 I (Final Project Presentation I)

17  2020/06/22  期末報告 II (Final Project Presentation II)

18  2020/06/29  教師彈性補充教學
Question Answering and Dialogue Systems
Outline

• Question Answering

• Dialogue Systems
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 23th 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

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
myday@mail.tku.edu.tw

Chun Tu

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

Min-Yuh Day
Cheng-Chia Tsai
Wei-Chun Chung
Hsiu-Yuan Chang
Tzu-Jui Sun
Yuan-Jie Tsai
Jin-Kun Lin
Cheng-Hung Lee
Yu-Ming Guo
Yue-Da Lin
Wei-Ming Chen
Yun-Da Tsai
Cheng-Jhiih Han
Yi-Jing Lin
Yi-Heng Chiang
Ching-Yuan Chien

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

Wanchu Huang  Shi-Ya Zheng  I-Hsuan Huang  Tz-Rung Chen  Min-Chun Kuo  Yue-Da Lin  Yi-Jing Lin
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  Jhih-Yi Chen  Yu-Ling Kuo  Jian-Ting Lin

myday@mail.tku.edu.tw

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
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

Dialogue Intention Detection

AIML Dialogue Engine

Real Time Dialogue API

System Response Generator

RNN LSTM GRU

Deep Learning
TensorFlow

Python NLTK

Dialogue KB

IR

Cloud Resource

AIML KB

Document Retrieval

Answer Extraction

Answer Generation

Answer Validation

Answer
IMTKU Emotional Dialogue System Architecture

1. Retrieval-Based Model
2. Generation-Based Model
3. Emotion Classification Model
4. Response Ranking

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

Retrieval-Based Model

1. Post
2. Word Segmentation
3. Keywords
4. Boolean Query
5. Solr Matching
6. Building Index
7. Distinct Result Data
8. Emotion Matching
9. Emotion Classification
10. Word2Vec Similarity Ranking
11. Retrieval-Based Response
The system architecture of IMTKU generation-based model for NTCIR-14 STC-3

**Generation-Based Model**

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

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

STC3 Corpus → Chinese Segmentation using Jieba → Stop Words Removal → Word2Vec → 1.2 million data (300 dimensions) → 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><strong>NTCIR-12 STC-1</strong>&lt;br&gt;22 active participants</td>
<td>Twitter, Retrieval</td>
<td>Weibo, Retrieval</td>
<td></td>
</tr>
<tr>
<td><strong>NTCIR-13 STC-2</strong>&lt;br&gt;27 active participants</td>
<td>Yahoo! News, Retrieval+ Generation</td>
<td>Weibo, Retrieval+ Generation</td>
<td></td>
</tr>
<tr>
<td><strong>NTCIR-14 STC-3</strong></td>
<td>Chinese Emotional Conversation Generation (CECG) subtask</td>
<td>Multi-turn, task-oriented (helpdesk)</td>
<td>Weibo+English translations, distribution estimation for subjective annotations</td>
</tr>
</tbody>
</table>
# 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>Platform</td>
<td>PC</td>
<td>Web</td>
<td>Smartphone</td>
<td>Messaging</td>
</tr>
<tr>
<td></td>
<td>Desktop DOS, Windows, Mac OS</td>
<td>Browser Mosaic, Explorer, Chrome</td>
<td>Mobile OS iOS, Android</td>
<td>Messaging Apps WhatsApp, Messenger, Slack</td>
</tr>
<tr>
<td>Applications</td>
<td>Clients Excel, PPT, Lotus</td>
<td>Website Yahoo, Amazon</td>
<td>Apps Angry Birds, Instagram</td>
<td>Bots Weather, Travel</td>
</tr>
<tr>
<td></td>
<td>Native Screens</td>
<td>Web Pages</td>
<td>Native Mobile Screens</td>
<td>Message</td>
</tr>
<tr>
<td>UI/UX</td>
<td>Client-side</td>
<td>Server-side</td>
<td>Client-side</td>
<td>Server-side</td>
</tr>
<tr>
<td>S/w Dev</td>
<td>Client-side</td>
<td>Server-side</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: https://bbvaopen4u.com/en/actualidad/want-know-how-build-conversational-chatbot-here-are-some-tools
AI Dialogue
System
Dialogue Subtasks

Dialogue subtasks

Dialogue Generation
- 9 leaderboards
- 35 papers with code

Dialogue State Tracking
- 2 leaderboards
- 30 papers with code

Visual Dialog
- 3 leaderboards
- 28 papers with code

Task-Oriented Dialogue Systems
- 20 papers with code

Goal-Oriented Dialog
- 15 papers with code

Short-Text Conversation
- 6 papers with code

Goal-Oriented Dialogue Systems
- 5 papers with code

Task-Completion Dialogue Policy Learning
- 2 papers with code

Source: https://paperswithcode.com/area/natural-language-processing/dialogue
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

- **Impossible**
  - Open Domain
  - Retrieval-Based

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

- **General AI [Hardest]**
  - Open Domain
  - Generative-Based

- **Smart Machine [Hard]**
  - Closed Domain
  - Retrieval-Based

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

Hotel Chatbot

**Intent Detection**

- **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

**Slot Filling**

- **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

- 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 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

Message platforms

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
Bots Landscape

Bots with traction

Connectors/Shared Services

AI Tools: Natural Language Processing, Machine Learning, Speech & Voice Recognition

Bot Discovery

Bot developer frameworks and tools

Analytics

Messaging

How to Build Chatbots

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>✗</td>
<td>✔️</td>
</tr>
<tr>
<td>NLP support</td>
<td>but possible to integrate with middlewares</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>

## Comparison of Most Prominent AI Services

<table>
<thead>
<tr>
<th></th>
<th>wit.ai</th>
<th>api.ai</th>
<th>LUIS.ai</th>
<th>IBM Watson</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Free of charge</strong></td>
<td>![checkmark]</td>
<td>![checkmark] but has paid enterprise version</td>
<td>![checkmark] it is in beta and has transaction limits</td>
<td>30 days trial then priced for enterprise use</td>
</tr>
<tr>
<td><strong>Text and Speech processing</strong></td>
<td>![checkmark]</td>
<td>![checkmark]</td>
<td>![checkmark] with use of Cortana</td>
<td>![checkmark]</td>
</tr>
<tr>
<td><strong>Machine Learning Modeling</strong></td>
<td>![checkmark]</td>
<td>![checkmark]</td>
<td>![checkmark]</td>
<td>![checkmark]</td>
</tr>
<tr>
<td><strong>Support for Intents, Entities, Actions</strong></td>
<td>![checkmark] Intents used as trait entities, actions are combined operations</td>
<td>![checkmark] Intents is the main prediction mechanism. Domains of entities, intents and actions</td>
<td>![checkmark]</td>
<td>![checkmark]</td>
</tr>
<tr>
<td><strong>Pre-build entities for easy parsing of numbers, temperature, date, etc.</strong></td>
<td>![checkmark]</td>
<td>![checkmark]</td>
<td>![checkmark]</td>
<td>![checkmark]</td>
</tr>
<tr>
<td><strong>Integration to messaging platforms</strong></td>
<td>![cross] web service API</td>
<td>![checkmark] also has facility for deploying to heroku. Paid environment</td>
<td>![checkmark] integrated to Azure</td>
<td>![checkmark] possible via API</td>
</tr>
<tr>
<td><strong>Support of SDKs</strong></td>
<td>![checkmark] includes SDKs for Python, Node.js, Rust, C, Ruby, iOS, Android, Windows Phone</td>
<td>![checkmark] C#, Xamarin, Python, Node.js, iOS, Android, Windows Phone</td>
<td>![checkmark] enables building with Web Service API, Microsoft Bot Framework integration</td>
<td>![checkmark] Proprietary language “AlchemyLanguage”</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)

Fine-tuning BERT on Dialogue Slot Filling (SF)

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

Pre-trained Language Model (PLM)

Semi-supervised Sequence Learning
context2Vec
Pre-trained seq2seq

ULMFiT  ELMo
Multi-lingual  Transformer

MultiFiT
Cross-lingual

BERT

GPT
Bidirectional LM
Larger model
More data

GPT-2
Defense
Whole Word Masking

Grover

XLM
UDify
MT-DNN
MASS
UniLM
SpanBERT
RoBERTa
XLNet

ERNIE
(Tsinghua)
Neural entity linker

KnowBert

VideoBERT
CBT
ViLBERT
VisualBERT
B2T2
Unicoder-VL
LXMER
VL-BERT
UNITER

ERNIE (Baidu)
BERT-wwm

By Xiaozhi Wang & Zhengyan Zhang @THUNLP

Source: https://github.com/thunlp/PLMpapers
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://summaris.es/">https://summaris.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/nyu-dl/SearchQA">https://github.com/nyu-dl/SearchQA</a></td>
</tr>
<tr>
<td></td>
<td>SearchQA</td>
<td></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>

Dialogue on Airline Travel Information System (ATIS)
The ATIS (Airline Travel Information System) Dataset

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: State-of-the-art Methods

<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](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
Restaurants Dialogue Datasets

- MIT Restaurant Corpus
  - [https://groups.csail.mit.edu/sls/downloads/restaurant/](https://groups.csail.mit.edu/sls/downloads/restaurant/)

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

- DSTC2 (Dialog State Tracking Challenge 2 & 3)
  - [http://camdial.org/~mh521/dstc/](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>
</table>
| **usr:** I’m looking for a college type attraction.  
... | **usr:** 你好，可以帮我推荐一个评分是4.5分以上的景点吗？  
Hello, could you recommend an attraction with a rating of 4.5 or higher? |
| **usr:** I would like to visit in town centre please.  
... | **sys:** 天安门城楼，簋街小吃和北京欢乐谷都是很不错的地方呢。  
Tiananmen, Gui Street, and Beijing Happy Valley are very nice places. |
| **usr:** Can you find an Indian restaurant for me that is also in the town centre? | **usr:** 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?  
I like Beijing Happy Valley. What hotels are around this attraction? |
| **Schema** | **sys:** 那可多了，有A酒店, B酒店, C酒店。  
There are many, such as hotel A, hotel B, and hotel C. |
| **usr:** I want a hotel in San Diego and I want to check out on Thursday next week.  
... | **usr:** 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?  
Great! I am planning to find a hotel to stay near the attraction. Which one has a rating of 4 or higher and offers wake-up call service? |
| **usr:** I need a one way flight to go there. | **sys:** ... |
# 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=, plate number=


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

Python in Google Colab (Python101)

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

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/imtkupython101
Python in Google Colab (Python101)

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

Table of contents

RandomizedSearchCV
Sentiment Analysis
  - Unsupervised
  - Supervised
  - 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

+ Code + Text

<table>
<thead>
<tr>
<th>Layer (type)</th>
<th>Output Shape</th>
<th>Param #</th>
<th>Connected to</th>
</tr>
</thead>
<tbody>
<tr>
<td>input_1 (InputLayer)</td>
<td>([None, 384])</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>input_3 (InputLayer)</td>
<td>([None, 384])</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>input_2 (InputLayer)</td>
<td>([None, 384])</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>tf_bert_model (TFBertModel)</td>
<td>([None, 384, 768]), (109482240)</td>
<td>0</td>
<td>input_1[0][0]</td>
</tr>
<tr>
<td>start_logit (Dense)</td>
<td>(None, 384, 1)</td>
<td>768</td>
<td>tf_bert_model[0][0]</td>
</tr>
<tr>
<td>end_logit (Dense)</td>
<td>(None, 384, 1)</td>
<td>768</td>
<td>tf_bert_model[0][0]</td>
</tr>
<tr>
<td>flatten (Flatten)</td>
<td>(None, 384)</td>
<td>0</td>
<td>start_logit[0][0]</td>
</tr>
<tr>
<td>flatten_1 (Flatten)</td>
<td>(None, 384)</td>
<td>0</td>
<td>end_logit[0][0]</td>
</tr>
<tr>
<td>activation_7 (Activation)</td>
<td>(None, 384)</td>
<td>0</td>
<td>flatten[0][0]</td>
</tr>
<tr>
<td>activation_8 (Activation)</td>
<td>(None, 384)</td>
<td>0</td>
<td>flatten_1[0][0]</td>
</tr>
</tbody>
</table>

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/imtkupython101
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)
{
   '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/imtkupython101
Python in Google Colab (Python101)

https://colab.research.google.com/drive/1FEG6DnGvfwUbeo4zJ1zTunjMqf2RkCrT

https://tinyurl.com/imtkupython101
Python in Google Colab (Python101)

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

```python
# Naive NER 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)
    active_slot_words = [word]
    active_slot_name = new_slot_name

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)
```

https://tinyurl.com/imtkupython101
# 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://summarize.es/">https://summarize.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 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>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>RACE</td>
<td><a href="http://www.qizhixie.com/data/RACE_leaderboard">http://www.qizhixie.com/data/RACE_leaderboard</a></td>
</tr>
<tr>
<td></td>
<td>SQuAD</td>
<td><a href="https://rajpurkar.github.io/SQuAD_leaderboard">https://rajpurkar.github.io/SQuAD_leaderboard</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/bdining/quasar">https://github.com/bdining/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/textsql-data/tree/master/data">https://github.com/jkkummerfeld/textsql-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>Sentiment Analysis</td>
<td>IMDB Reviews</td>
<td><a href="http://ai.stanford.edu/~amaas/data/sentiment/">http://ai.stanford.edu/~amaas/data/sentiment/</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>Subjectivity Dataset</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></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>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></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
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://m2dsupsdliclass.github.io/lectures-labs/