Natural Language Processing with Deep Learning

Introduction

- Class Plan
- NLP Introduction
- Class Project

Class Plan

Course Materials

- Materials from Stanford 224n (Chris Manning)

  https://web.stanford.edu/class/cs224n/

- Papers (in the separate doc)

- Speech and Language Processing 3rd Edition, Dan Jurafsky and James H. Martin

  https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf

Grading

- Participation (10%) :
  - Attendance
  - Interactions with professors during lectures and presentations and contributions during practical time are evaluated.
- Mid-term Exam (30%)
  - 8th week (Exact time will be announced separately)
- Project (40%)
  - Students work on team projects during the semester, replacing final grades. Details will be explained in the first hour
- Quiz (20%)

Course Overview

1. Word Embedding
2. Neural Networks
3. Dependency Parsing
4. Recurrent Neural Networks
5. Seq2Seq Model and Neural Machine Translation
6. Attention Mechanism
7. Transformer
8. Pretrained Language Models
9. Mid-term Exam
10. Question Answering
11. Representing and Using Knowledge in NLP
12. Chatbots and Dialog System
13. Natural Language Generation
14. Information Extraction
15. Project Report and Evaluation

Course Overview: 1st week

1. Word Embeddings Term Project Handout:

(1) Project Handout (Robust QA track)

(2) Project Handout (IID SQuAD track)

(3) End to End Question-Answering System Using NLP and SQuAD Dataset

Suggested Readings:

(1) Efficient Estimation of Word Representations in Vector Space (original word2vec paper)

(2) Distributed Representations of Words and Phrases and their Compositionality (negative sampling paper)

Additional Readings:

(1) GloVe: Global Vectors for Word Representation (original GloVe paper)

(2) Improving Distributional Similarity with Lessons Learned from Word Embeddings

(3) Evaluation methods for unsupervised word embeddings

(4) A Latent Variable Model Approach to PMI-based Word Embeddings

(5) Linear Algebraic Structure of Word Senses, with Applications to Polysemy

(6) On the Dimensionality of Word Embedding

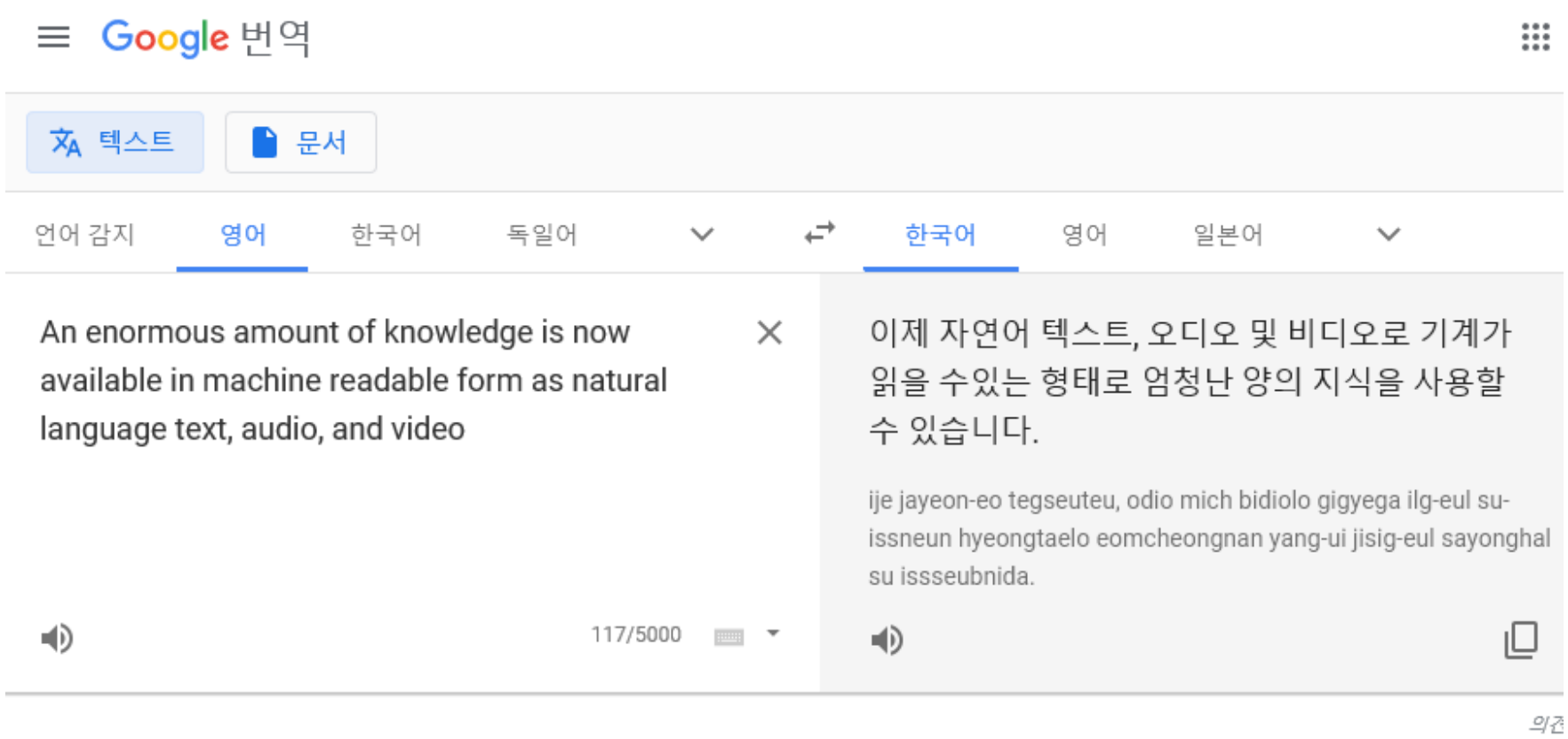
NLP Introduction

Artificial Intelligence: Core Areas

Class	Human function	Definition	Sub areas
Natural Language /Speech Processing		Supports machine-human dialogue by understanding the meaning of speech and text	Natural Language Processing, Conversational Systems, Q&A, Machine Translation, Speech Processing, Speech Recognition
Computer Vision		Understanding the visual information of the surrounding environment and meaning	Object Recognition/Tracking, Image Search, Human Recognition/Understanding, Scene Understanding, Spatial Information Understanding, Image Improvement
Data Analysis / Inference		Deriving new facts based on logic/probability	Inference, prediction, planning, recommendation
Robotics / Control		Physical control technology that simulates human-like movements	operation control, movement control
Machine Learning		Self-learning the basic characteristics of data	machine learning, deep learning, reinforcement learning

NLP Applications

Machine Translation



Question Answering

IBM Watson: Jeopardy! winner

http://www.geekwire.com/2013/ibm-takes-watson-cloud/

Google's question answering system

Google

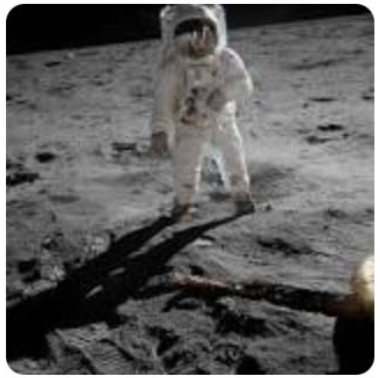
when did the apollo 11 land on the moon

All Images News Videos Shopping More Settings Tools

About 5,020,000 results (0.85 seconds)

Apollo 11 / Land date

July 24, 1969



People also search for

Apollo 13
April 17, 1970

Apollo 12
November 24, 1969

Apollo 1
March 7, 1967

Personal Assistant

https://reachbyte.com/siri-cortana-alexa-google-assistant-bixby/

NLP: Requirement

- Effective handling of **different characteristics of each language**
- **Common processing methods** for different languages, genres, styles and forms
- Computational **efficiency** at build time and runtime
- Strong theoretical guarantee (e.g. **convergence**)
- **High accuracy** when judged by an expert

Class Project

Project: Question Answering

- Input: a paragraph and a question about that paragraph
- Output: answer

1. Building a QA system for the SQuAD dataset
2. **Robust QA**
   - unseen question answering datasets, along with a small training set of 128 examples for additional finetuning.
     - Note: you are allowed to use only DistilBERT [8] as the pre-trained transformer model

SQuAD

Question: Why was Tesla returned to Gospic?
Context paragraph: On 24 March 1879, Tesla was returned to Gospic under police guard for not having a residence permit. On 17 April 1879, Milutin Tesla died at the age of 60 after contracting an unspecified illness (although some sources say that he died of a stroke). During that year, Tesla taught a large class of students in his old school, Higher Real Gymnasium, in Gospic.
Answer: not having a residence permit

- Paragraphs in SQuAD are from Wikipedia
- Questions and answers were crowdsourced using Amazon Mechanical Turk
- SQuAD 2.0: some questions cannot be answered using the paragraph
- https://rajpurkar.github.io/SQuAD-explorer/explore/v2.0/dev/.

SQuAD : Answer Evaluation

- Every answerable SQuAD question has three answers provided
- Metrics
  - Exact Match (EM):
    - E.g. When the ground truth is 'Albert Einstein', EM for 'Einstein' is score is 0
  - F1 - harmonic mean of precision and recall
    - E.g. 100% precision and 50% recall →

      2×prediction×recall/(precision+recall) = 2*50*100/(100+50) = 66.67%.

Robust QA

- Every answerable SQuAD question has three answers provided
- Code provided: preprocessing the data and computing the evaluation metrics, and to train a fully-functional neural baseline
- You're not required to implement something original
- The best projects will pursue some originality with improvements over the baseline

Datasets & GitHub

Dataset	Question Source	Passage Source	Train	dev	Test
		in-domain datasets			
SQuAD [5]	Crowdsourced	Wikipedia	50000	10,507	-
NewsQA [7]	Crowdsourced	News articles	50000	4,212	-
Natural Questions [6]	Search logs	Wikipedia	50000	12,836	-
		oo-domain datasets			
DuoRC [9]	Crowdsourced	Movie reviews	127	126	1248
RACE [10]	Teachers	Examinations	127	128	419
RelationExtraction [11]	Synthetic	Wikipedia	127	128	2693

https://github.com/michiyasunaga/robustqa

Evaluation

- the creativity, complexity and technical correctness of your approach
- your thoroughness in exploring and comparing various approaches, the strength of your results,
- the effort you applied, and the quality of your write-up, evaluation, and error analysis
- Implementing a small number of things with good results and thorough experimentation/analysis is better than implementing a large number of things that don't work

Evaluation Policy

Subject	Evaluation items	Scores
Report writing	1. Is the project report well organized and written in an easy-to-under stand manner ?	**15**
	2. Are citations appropriate ?	**10**
Achievement	1. What is the contribution of the work ?	**10**
	2. How much has it improved in terms of performance?	**15**
Originality	1. What is the originality in the research content and how much did it affect the performance improvement of the algorithm?	**20**
	2. Which part of the source codes are written by authors ? Are those codes are critical to the the performance improvement of the algorith m?	**20**
Presentation	1. Is the presentation well prepared?	**5**
	2. Does the presenter explain well simply and clearly?	**5**
Total		**100**