Building Large Language Model Applications

Introduction to Natural Language Processing

Hamza Farooq Dr. Saima Hassan







Learning outcomes

- NLP Overview
- History of NLP
- Applications of NLP
- Conclusion

NEW YORK NEWS Articles for you **Natural Language Processing**

We live in a world of NLP



What is NLP from the experts' perspective?

"Natural language is the most important part of artificial intelligence." John Searle

"Natural language processing is a cornerstone of artificial intelligence, allowing computers to read and understand human language, as well as to produce and recognize speech." Ginni Rometty

"Natural language processing is one of the most important fields in artificial intelligence and also one of the most difficult." Dan Jurafsky



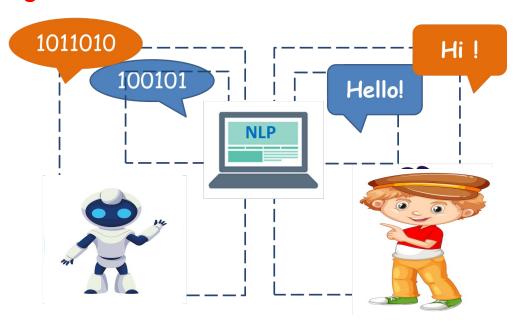






What is NLP?

- NLP process information contained in a natural language text
- It deals with the interaction between human and computer using natural language







```
Read and convert alphabetic imput cards (Russian sentences)
Initialize complete (left-balf partial) word look-up subprogram
Initialize lat Emglish inquase-block cell
Initialize lat diacritic block cell
                                     Initialize 1st cell of Russian input word
                                                                                                                   Test next input word
                                                                                                                   cell for 0
                                    Extraxt 1st character of imput word 1st cell
setting "Thumb index" drum address, Sp
                                                                                                                      (No)
                                                                                                                                    (Yes)
                                     and drum location, ∞r
                                    Table look up for alpha region index (8p , 00 p)
                                                                                                                  cell for 0
                                     Copy first cell of dictionary word and test for '+'
                                                                                                           Set discritic loop exit test
  word look-up
   Copy all succeeding cells of dictionary line associated
                                                                                           word and
  with successfully tested Russian word
 Initialize dictionary word 1st comparison cell
Initialize input word 1st cell
                                                                                           Copy English No. 1, English No. 2, and
  Initialize for k equal to 1st character extraction and
                                                                                           discritic associated with dictionary
  Extract k-th character (15 k55) of input word cell, and
                                                                                           Test whether input word was complete or
left-half portial (hyphen search)
 kth character of imput word cell is not equal to 0000000:
                                                                                               Input word was complete: modify impword, English block copy cell, and discritic block copy cell
  word corresponding
  If difference equals 0000000, there is character agreement,
                                                                          Input word was left-half partial beginning with
input word cell containing byphen: shift all re-
maining cells left, so that 1st character following
Yest: is k+1 less than or equal to 5?
                                                                           hyphen occupies 1st character position. Set
address of first cell of mext input word
 Nedify imput word cell address
Nedify dictionary word cell address
                                                                                                     Set last partial word cell to zero
  If difference is not equal to 0000000, there
                                                                           k-th character of dictionary word cell is a hyphen:
Initialize table look-up drum address and location,
  is no character agreement, so test k-th cha-
 racter of dictionary word cell for existence of hyphen
                                                                          using k-th character of input word cell as argument.
Set address parts of instructions in right-half
partial word shifting subroutine. Store drum address
 k-th character of dictionary word cell is not
a hypeen: hence, there is no alphabetical
                                                                          and location of English No. 1 (belonging to left-half
partial word) in English and discritic block copy
 agreement in this character position
                                                                          routine
                                                                                        Copy 1st cell of next dictionary
                                "copy annihilate" succeeding cells
                               of dictionary word line
              Test new 1st character against old
                                                                                         Skip to next dictionary word and ...
                         lst character
                                                                                                                             (No agreement
                          Initialize partial word 1st cell
                                                                                                 Copy 1st cell of dictionary, item and test for '+
                         Modify-initialize English block copy cells
Modify-initialize discritic block copy cells
 Copy all succeeding cells of dictionary word line
Initialize for dictionary word 1st comparison cell
                                                                                   Skip to mext dictionary word ...
 Initialize for character extraction | Extract k-th character of partial k-th character is not equal and comparison; k equals l | word cell and test for 0000000 to 00000000; compare k-th
  No agreement: "copy annihilate" to
next dictionary word, and ...
                                                                                                                 characters of partial word
                                                        Agreement: modify k into k+1 characters of partia and dictionary word is k+1 less than or equal to 577
 k-th character is equal to DUUDDDU: Copy
English No. 1 and English No. 2, and asso-
ciated discritic numbers. Modify English
                                                                                                                             Modify partial word
                                                                        Operational syntax program
   and discritic block copy cells and
```

```
Welcome to
                    EEEEEE LL
                                                  AA AA
                                                  AAAAAAA
                    EEEEEE LLLLLL IIII ZZZZZZ
  Eliza is a mock Rogerian psychotherapist.
  The original program was described by Joseph Weizenbaum in 1966.
  This implementation by Norbert Landsteiner 2005.
ELIZA: Is something troubling you?
      Men are all alike.
ELIZA: What is the connection, do you suppose?
      They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
      Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
      He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU: It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
```

https://en.wikipedia.org/wiki/ELIZA

1960s

Gemini ▼

```
(defrule 52

if (site culture is blood)
  (gram organism is neg)
  (morphology organism is rod)
  (burn patient is serious)
  then .4
  (identity organism is pseudomonas))
Rule 52:

If
  1) THE SITE OF THE CULTURE IS BLOOD
  2) THE GRAM OF THE ORGANISM IS NEG
  3) THE MORPHOLOGY OF THE ORGANISM IS ROD
  4) THE BURN OF THE PATIENT IS SERIOUS
Then there is weakly suggestive evidence (0.4) that
  1) THE IDENTITY OF THE ORGANISM IS PSEUDOMONAS
```

1970s and 80s

write a poem on AI in four lines

Here's a poem on AI in four lines:

A mind of metal, thoughts take flight, Learning fast, both day and night. Solving problems, complex and grand, Al's future, close at hand.

. . .

1950s

LOOKUP PROCEDURES.

A FLOWCHART OF PART OF THE IBM'S DICTIONARY

Today



Go beyond the keyword matching





- Identify the structure and meaning of words, sentences, texts and conversations
- Deep understanding of broad language
- NLP is all around us



Knowledge requirements for Machine

Phonetics and Phonology: Knowledge about linguistic sounds

Morphology: Knowledge of the meaningful components of words

Syntax: Knowledge of the structural relationships between words

Semantics: Knowledge of meaning

Pragmatics: Knowledge of the relationship of meaning to the goals and intentions of the speaker

Discourse: Knowledge about linguistic units larger than a single utterance







What are the NLP applications you interact with daily?

- Search engines
- Google Translate
- Social Media
- Job Seeking
- ...







Most ideas stem from Academia, but big guys have (several) strong NLP research labs (like Microsoft, Yahoo, AT&T, IBM, etc.)



Applications

Conversational agents contains:

- Speech recognition
- Language analysis
- Dialogue processing
- Information retrieval
- Text to speech





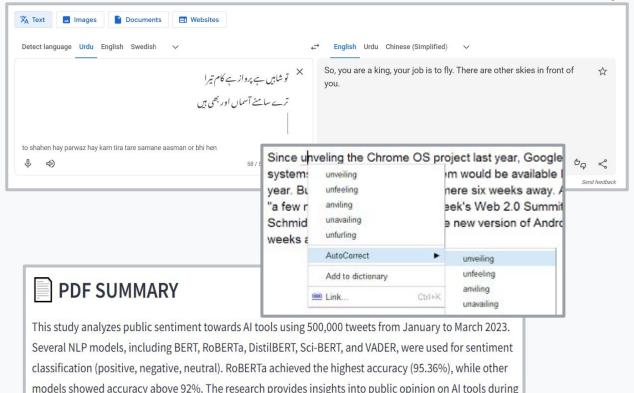


Applications

- Machine Translation
- Summarization
- Auto Completion
- Spell Correction

this period.

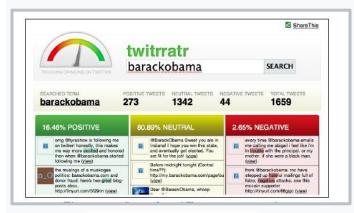




Applications

- Sentiment Analysis
- Text Classification

Many More...





DOI: 10.13140/RG.2.1.1809.1044







Conclusion

What is NLP?

Key challenges and applications