Name: Megha Nagabhushan

ClassID: 15.

1. Given the text example below, show the following NLP operations.

    a. Sentence detection/Tokenization

    b. Lemmatization

    c. POS tagging/Chunking

    d. Parsing

    e. Name Entity Recognition

    f. Co-reference Resolution

2. Summarize and draw a knowledge graph.

CHICAGO (AP) — Citing high fuel prices, United Airlines said Friday it has increased fares by $6 per round trip on flights to some cities also served by lower-cost carriers. American Airlines, a unit AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

---

① Natural Language processing refers to the use and ability of systems to process sentences in a natural language such as English, rather than in a specialized artificial computer language such as C++.

It involves.

a) Tokenization : tokenizing text into words.

b) Lemmatization / Stemming:

Stemming is the process of reducing words to their stem, base or root form - generally a written word

. Lemmatization is the process of grouping together the different inflected forms of a word so they can be analysed as a single term.

c) POS Tagging / Chunking

Parts - of - Speech Tagging. Uses a pre-defined Tagset for this process.

d) Parsing

It is the process of analyzing a sentence by taking each word and determining its structure from its constituent parts. It uses two components : a parses and a grammar.

e) Name Entity Recognition.

It is a subtask of Named-Entity-Recognition. Information extraction that seeks to locate and classify named entities in text into pre-defined categories such as names of persons, organizations, locations, expressions of times, quantities, monetary values, percentages etc.

f) Co-reference Resolution

It is the task of finding all expressions that refer to the same entity in a text. It is important step for higher NLP tasks like document summarization, question answering and information extraction.

annotation → lemmatization → POS → Named Entity
                              Tagging   Recognition

Airlines → Airlines → NNPS → ORGANIZATION

said → say → VBD

Friday → Friday → NNP → DATE

it → it → PRP

has → have → VBZ

increased → increase → VBN

fares → fare → NNS

by → by → IN

$ → $ → MONEY

6 → 6 → CD → MONEY

per → per → IN

round → round → NN

trip → trip → NN

on → on → IN

flights → flight → NNS

to → to → TO

some → some → DT

cities → city → NNS

also → also → RB

served → serve → VBD

by → by → IN

lower-cost → lower-cost → JJ

carriers → carriers → NNS

. → . → .

American → American → NNP

Airlines → Airlines → NNPS

, → , → ,

a → a → DT

unit → unit → NN

AMR → AMR → NNP

, → , → ,

immediately → immediately → RB

matched → match → VBD

the → the → DT

move → move → NN

, → , → ,

spokesman → spokesman → NN

Tim → Tim → NNP → PERSON

Wagner → Wagner → NNP → PERSON

said → say → VBD

. → . → .

United → United → NNP → ORGANIZATION

, → , → ,

a → a → DT

unit → unit → NN

of → of → IN

American → American → NNP → ORGANIZATION

Airlines → Airlines → NNPS → ORGANIZATION

UAL → UAL → NNP → ORGANIZATION

, → , → ,

said → say → VBD

the → the → DT

increase → increase → NN

took → take → VBD

effect → effect → NN

Thursday → Thursday → NNP → DATE.

night → night → NN → TIME

and → and → CC

applies → apply → VBZ

to → to → TO

most → most → JJS.

routes → route → NNS

where → where → WRB

it → it → PRP

competes → compete → VBZ

against → against → IN

discount → discount → NN

carriers → carrier → NNS

, → , → ,

such → such → JJ

as → as → IN

Chicago → Chicago → NNP → LOCATION.

Step 2: Information Retrieval and Extraction.

CHICAGO (AP) — Citing high fuel prices, United Airlines said Friday it has increased fares by $6 per trip on flights to some cities also served by lower-cost carriers. American Airlines, a unit AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

Step 3: Topic discovery.

Organization	People	Places
United Airlines American Airlines AMR UAL	Tim Wagner.	Chicago Dallas Atlanta Denver San Francisco Los Angeles New York

Step: 4   Knowledge Graph Constuction.

AMR — unit of → American Airlines

American Airlines — matched the price hike of → United Airlines

Tim Wagner — spokesperson of → American Airlines

UAL — unit of → United Airlines

United Airlines — increased fares by $6 in these routes due to high fuel prices → Chicago

Chicago → Dallas

Chicago → Atlanta

Denver → San Francisco

Los Angeles → New York