CS5560 Knowledge Discovery and Management

Problem Set (PS-2)

Aug 28, 2018

Name: Lakshmi Korrapati

ClassID: 4

- Given the text example below, show the following NLP operations.
    - Sentence detection/Tokenization
    - Lemmatization
    - POS tagging/Chunking
    - Parsing
    - Name Entity Recognition
    - Co-reference Resolution
- 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.

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

a.Sentence detection/Tokenization
Ans)
CHICAGO-1, -LRB—2, AP-3, -RRB—4, ---5, Citing-6, high-7, fuel-8, prices-9, ,-10, United-11, Airlines-12, said-13, Friday-14, it-15, has-16, increased-17, fares-18, by-19, $-20, 6-21, per-22, round-23, trip-24, on-25, flights-26, to-27, some-28, cities-29, also-30, served-31, by-32, lower-cost-33, carriers-34, .-35, American-1, Airlines-2, ,-3, a-4, unit-5, AMR-6, ,-7, immediately-8, matched-9, the-10, move-11, ,-12, spokesman-13, Tim-14, Wagner-15, said-16, .-17, United-1, ,-2, a-3, unit-4, of-5, UAL-6 ,-7, said-8, the-9, increase-10, took-11, effect-12, Thursday-13, night-14, and-15, applies-16, to-17, most-18, routes-19, where-20, it-21, competes-22, against-23, discount-24, carriers-25, ,-26, such-27, as-28, Chicago-29, to-30, Dallas-31, and-32, Atlanta-33, and-34, Denver-35, to-36, San-37, Francisco-38, ,-39, Los-40, Angeles-41, and-42, New-43, York-44, .-45

b.Lemmatization
Ans)
CHICAGO, -lrb-, AP, -rrb-, --, cite, high, fuel, price, ,, United, Airlines, say, Friday, it, have, increase, fare, by, $, 6, per, round, trip, on, flight, to, some, city, also, serve, by, lower-cost, carrier, ., American, Airlines, ,, a, unit, AMR, ,, immediately, match, the, move, ,, spokesman, Tim, Wagner, say, ., United, ,, a, unit, of, UAL, ,, say, the, increase, take, effect, Thursday, night, and, apply, to, most, route, where, it, compete, against, discount, carrier, ,, such, as, Chicago, to, Dallas, and, Atlanta, and, Denver, to, San, Francisco, ,, Los, Angeles, and, New, York, .

c.POS tagging/Chunking
CHICAGO-1:NNP, -LRB--2:-LRB-, AP-3:NNP, -RRB--4:-RRB-, ---5::, Citing-6:VBG, high-7:JJ, fuel-8:NN, prices-9:NNS, ,-10::, United-11:NNP, Airlines-12:NNPS, said-13:VBD, Friday-14:NNP, it-15:PRP, has-16:VBZ, increased-17:VBN, fares-18:NNS, by-19:IN, $-20:$,

6-21:CD, per-22:IN, round-23:NN, trip-24:NN, on-25:IN, flights-26:NNS, to-27:TO, some-28:DT, cities-29:NNS, also-30:RB, served-31:VBD, by-32:IN, lower-cost-33:JJ, carriers-34:NNS, .-35:., American-1:NNP, Airlines-2:NNPS, ,-3:,, a-4:DT, unit-5:NN, AMR, 6:NNP, ,-7:,, immediately-8:RB, matched-9:VBD, the-10:DT, move-11:NN, ,-12:,, spokesman-13:NN, Tim-14:NNP, Wagner-15:NNP, said-16:VBD, .-17:., United-1:NNP, ,-2:,, a-3:DT, unit-4:NN, of-5:IN, UAL-6:NNP, ,-7:,, said-8:VBD, the-9:DT, increase-10:NN, took-11:VBD, effect-12:NN, Thursday-13:NNP, night-14:NN, and-15:CC, applies-16:VBZ, to-17:TO, most-18:JJS, routes-19:NNS, where-20:WRB, it-21:PRP, competes-22:VBZ, against-23:IN, discount-24:NN, carriers-25:NNS, ,-26:,, such-27:JJ, as-28:IN, Chicago-29:NNP, to-30:TO, Dallas-31:NNP, and-32:CC, Atlanta-33:NNP, and-34:CC, Denver-35:NNP, to-36:TO, San-37:NNP, Francisco-38:NNP, ,-39:,, Los-40:NNP, Angeles-41:NNP, and-42:CC, New-43:NNP, York-44:NNP,.-45:.

d.Parsing
Ans)
CHICAGO/NNP LRB/LRB AP/NNP RRB/RRB -/ Citing/VBG high/JJ fuel/NN prices/NNS ,/, United/NNP Airlines/NNPS said/VBD Friday/NNP it/PRP has/VBZ increased/VBN fares/NNS by/IN $/$ 6/CD per/IN round/NN trip/NN on/IN flights/NNS to/TO some/DT cities/NNS also/ RB served/VBD by/IN lower-cost/JJ carriers/NNS ./. American/NNP Airlines/NNPS ,/, a-/DT unit/NN AMR/NNP ,/, immediately/RB matched/VBD the/DT move/NN ,/, spokesman/NN Tim/ NNP Wagner/NNP said/VBD ./. United/NNP ,/, a/DT unit/NN of/IN UAL/NNP ,/, said/VBD the/ DT increase/NN took/VBD effect/NN Thursday/NNP night/NN and/CC applies/VBZ to/TO most/JJS routes/NNS where/WRB it/PRP competes/VBZ, against/IN, discount/NN carriers/ NNS ,/, such/JJ as/IN Chicago/NNP to/TO Dallas/NNP and/CC Atlanta/NNP and/CC Denver/ NNP to/TO San/NNP Francisco/NNP ,/, Los/NNP Angeles/NNP and/CC New/NNP YorkNNP./.

e.Name Entity Recognition
Ans)
United- Organization
Airlines- Organization
Friday- Date
$- Money
6- Money
American- Organization
Airlines- Organization
AMR- Organization
Tim- Person
Wagner- Person
United- Organization
UAL- Organization
Thursday- Date
Chicago- Location
Dallas- Location
Atlanta- Location

Denver- Location
San- Location
Francisco- Location
Los- Location
Angeles- Location
New- Location
York- Location
Remaining all are null.

f.Co-reference Resolution
Ans)

CHAIN1-["United Airlines" in sentence 1, "it" in sentence 1], CHAIN2-["high fuel prices" in sentence 1], CHAIN3-["Friday" in sentence 1], CHAIN5-["fares" in sentence 1], CHAIN6-["round trip" in sentence 1], CHAIN7-["flights" in sentence 1], CHAIN8-["some cities also served by lower-cost carriers" in sentence 1], CHAIN9-["lower-cost carriers" in sentence 1], CHAIN10-["American Airlines" in sentence 2, "a unit AMR" in sentence 2], CHAIN11-["AMR" in sentence 2], CHAIN12-["American Airlines , a unit AMR" in sentence 2], CHAIN14-["the move" in sentence 2], CHAIN15-["spokesman Tim Wagner" in sentence 2], CHAIN16-["United" in sentence 3, "a unit of UAL" in sentence 3], CHAIN17-["UAL" in sentence 3], CHAIN18-["Dallas" in sentence 3], CHAIN19-["Atlanta" in sentence 3], CHAIN20-["San Francisco" in sentence 3], CHAIN21-["Los Angeles" in sentence 3], CHAIN22-["New York" in sentence 3], CHAIN23-["United , a unit of UAL" in sentence 3], CHAIN25-["the increase" in sentence 3, "it" in sentence 3], CHAIN26-["Thursday night" in sentence 3], CHAIN27-["most routes" in sentence 3], CHAIN29-["discount carriers , such as Chicago to Dallas and Atlanta and Denver to San Francisco , Los Angeles and New York" in sentence 3], CHAIN30-["discount carriers" in sentence 3], CHAIN31-["Chicago to Dallas and Atlanta and Denver to San Francisco , Los Angeles and New York" in sentence 3], CHAIN32-["Chicago to Dallas and Atlanta" in sentence 3], CHAIN33-["Dallas and Atlanta" in sentence 3], CHAIN34-["Denver to San Francisco , Los Angeles and New York" in sentence 3], CHAIN35-["San Francisco , Los Angeles and New York" in sentence 3]

2.Summarize and draw a knowledge graph.
Ans)