

Student ID _____ :

Name-Surname _____ :

Natural Language Processing final exam (2018-2019 Fall)

(Please use free space for draft and fit your answer to boxes.)

1. Fill in the blanks with the most appropriate terms. Each question is five points (5P).

a. Levenshtein Distance is using for measuring for _____.

b. Vector Space Models biggest disadvantage is _____.

c. Difference between Lexical Similarity and Semantic Similarity is _____.

d. There are a lots of WordNet relations. Two of them are _____.

e. _____ level semantic similarity deal with the conceptual part of a word.

f. Word level similarity approaches are grouped in two titles: _____ and _____.

g. In Wordnet based similarity methods, the nodes in the net represent the _____ and the edges are the _____.

h. _____ is an algorithm considering a user who clicks on hyperlinks in the web pages.

i. Dialogue Systems can be classified as _____ and _____.

j. _____ is one of the Machine Translation approaches.

k. The difference between keyword and keyphrase is _____.

l. There are two approaches to document summarization: _____ and _____.

m. One of the paraphrase types is _____.

n. CLOM is the acronym of _____.

2. (30P) We want to prepare a Turkish WordNet by using a corpus. Giving your comments, explain your solution (How can we do it? What do we need?)