ASSIGNMENT 1 - MACHINE LEARNING IN COMPUTATIONAL LINGUISTICS

1) **Supervised Machine Learning**: The data that we have, consists of labels. The purpose of these labels is that when via a training method, we make predictions, some of these predictions can go wrong. Thus, with the help of labels, we can correct the wrongly classified items. Examples are: Classification and regression. Detecting whether the mail is spam/not-spam. Predicting housing price when we are given the data including the price and square feet of the house.
Unsupervised Machine Learning: When the data we have, is not labeled and it does not have a known result. So with the help of the model, we try to deduce general rules to organize the data, like finding similarity amongst the data (clustering).
Examples are: clustering, association rule mining.

Eager Learning: In eager learning, we generalize before seeing the query (or abstract over data). Example: Support vector machines, neural networks.
Lazy Learning: Generalization is delayed until a query is made to the system. (Or do not abstract over data). Example: Instance based learning.

Descriptive Model: In a descriptive model, there is a mathematical model that describes certain events, and relationships that caused those events to happen. Example: if the cost of an item increases by a certain amount every month, we can predict what the cost would be, in the next month.
Predictive Model: There are no events or relationships between those events. It just means accurately being able to predict the future outcomes. Example: If you provide water with a cold temperature, it will change its state from liquid to solid and become ice.

2) POS Tagging is a classification problem because we classify the sentences into various parts of speech. When we feed our word to be processed by the classifier, we also use the surrounding tokens to determine what part of speech our word is, since several words can be two parts of speech.
The used features are: words, word n-grams, POS tags, POS n-grams, combination of words + POS tags, how far is the word from the beginning of the sentence, what is the syntactic context, is there a comma in the vicinity, how many punctuation signs are there.
It is not a good idea to use the whole sentence as a feature, because the order of the words does not matter. Also, a feature should be such that it helps improve accuracy. A whole sentence does not appear again and again, thus reducing accuracy.

3) Title: The Interaction of Knowledge Sources in Word Sense Disambiguation
Authors: Mark Stevenson (University of Sheffield), Yorick Wilks (University of Sheffield)
Year: 2001
Conference/Journal: Computational Linguistics, Volume 27, Number 3
Issue/Conference Venue: ISSN 0891-2017
Pages: Pages 321-349
ML Algorithm used: Supervised Learning

Title: A Fully Unsupervised Word Sense Disambiguation Method Using Dependency Knowledge
Authors: Ping Chen, Wei Ding, Chris Bowes, David Brown
Year: 2009
Conference/Journal: Computational Linguistics, Volume 35, Number 4
Issue/Conference Venue: Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the ACL
Pages: Pages 28-36
ML Algorithm used: Unsupervised Learning

Title: Word Sense Disambiguation Improves Information Retrieval
Authors: Zhi Zhong, Hwee Tou Ng
Year: 2012
Conference/Journal: Association for Computational Linguistics Stroudsburg, PA, USA ©2012
Issue/Conference Venue: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics
Pages: Pages 273-282
ML Algorithm used: MaxEnt

Title: A Machine Learning Approach to coreference Resolution of noun phrases
Authors: Wee Meng Soon, Hwee Tou Ng, Daniel Chung Yong Lim
Year: 2001
Conference/Journal: Volume 27, Number 4
Issue/Conference Venue: MIT Press Cambridge, MA, USA. ISSN 0891-2017
Pages: Pages 273-282
ML Algorithm used: Decision tree learning algorithm

Title: Machine Learning Approaches to Coreference Resolution
Authors: Nguy G. L.
Year: 2001
Issue/Conference Venue: ISBN 978-80-7378-065-4 © MATFYZPRESS
ML Algorithm used: Clustering, Decision tree learning algorithm

Title: Improving Machine Learning Approaches to Coreference Resolution
Authors: Vincent Ng and Claire Cardie
Year: 2002
Issue/Conference Venue: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)
Pages: 104-111
ML Algorithm used: Clustering

Title: The Importance of Syntactic Parsing and Inference in Semantic Role Labeling
Authors: Vasin Punyakanok, Dan Roth, Wen-tau Yih
Year: 2008
Conference/Journal: Volume 34 Issue 2, June 2008
Issue/Conference Venue: Journal: Computational Linguistics
Pages:257-287
ML Algorithm Used: Classification

Title: Semantic Role Labeling Using Support Vector Machines
Authors: Tomohiro Mitsumori, Masaki Murata, Yasushi Fukuda, Kouichi Doi, and Hirohumi Doi
Year: 2005
Issue/Conference Venue: CONLL '05 Proceedings of the Ninth Conference on Computational Natural Language Learning
Pages: 197-200
ML Algorithm Used: Support Vector Machines

Title: Calibrating Features For Semantic Role Labeling
Authors: Nianwen Xue, Martha Palmer
Year: 2004
Issue/Conference Venue: In Proceedings of EMNLP 2004
Pages: 88-94
ML Algorithm Used: Support Vector Machines

Title: Adapting Self-training for Semantic Role Labeling
Authors: Rasoul Samad Zadeh Kaljahi
Year: 2010
Issue/Conference Venue: Proceedings of the ACL 2010 Student Research Workshop
Pages: 91-96
ML Algorithm Used: Supervised Learning