Sinhgad College of Engineering

Department of Information Technology

2018-19 (Sem-II)

TE Seminar

Project Synopsis

|  |  |
| --- | --- |
| Group Number | A8 |
| Project Title | Word Sense Disambiguation (WSD) |
| Abstract | Word Sense Disambiguation (WSD) is the task of removing ambiguity in different senses of words. It is a core research field in computational linguistics dealing with the automatic assignment of senses to words occurring in a given context. Humans are inherently good at WSD and distinguish senses used in words through spoken language. Computers on the other hand have difficulties identifying correct senses of words. Various advancements have been made in the task of disambiguation using mainly four approaches: Knowledge-based, Supervised, Semi-Supervised, and Unsupervised. Better understanding of the human language will help computer's performance in various applications such as search engine optimization, information retrieval, information extraction, software assistants, and voice command interpretation. The objective of this work is to present a supervised neural network model using machine learning algorithms dedicated to the task of maximizing accuracy of sense detection. The input layer of the neural network will consist of nodes having binary values depending on the presence or absence of frequently occurring context words related to the ambiguous words. The output layer will consist of nodes equal to the number of senses the ambiguous word has. Training and testing of the model will be done using lexical resources such as SemCor or OMSTI. Accuracy will be calculated based on All- Word tasks from SemEval International Workshops. |
| Aim | The aim of this project is to create a model which maximizes accuracy of Word Sense Disambiguation (WSD) using Neural Networks. |
| Objectives | 1. To understand the problem concerning word sense ambiguity. 2. To study the previous research work done in the field of Word Sense Disambiguation. 3. To find and acquire necessary labelled datasets for the purpose of training and testing. 4. To design and train a WSD model based on machine learning and neural network algorithms, with the goal of maximizing accuracy. 5. To compute accuracy of the model using appropriate testing data, and to compare it with other WSD models already documented. |
| Software requirements | WordNet Lexical Database; Programming IDE; |
| Hardware requirements | - |