

NATURAL LANGUAGE PROCESSING

PROJECT

Project Title: Intelligent Matcher

Team Name: Matcher

Team Member: Praveen Erode Murugesan (Net ID: pem150030)

1. Problem Description:

1.1 Problem Statement:

The project aims at finding the appropriate service provider for a particular consumer where both of them expresses their interests in a common forum.

1.2 Explanation:

Let us consider a common forum where users express their interests and needs. The users are classified as consumer or a service provider based on their expressed statement. The service needed by a user is extracted from the statement and checked against all the service providers' statements, the ones that get matched are listed as results to that particular consumer.

Eg:

Consumer Statements

I need a cab

I want to buy carrots

Service Provider Statements

I sell vegetables

I am a taxi driver

*Note - matching statements are shown in same color

2. Proposed Solution and Implementation Details

2.1 Baseline System:

Base line system will be implemented for exactly matched statements. (i.e) *'I want to buy carrots'* will get matched only with Service Provider Statements that has *carrots* in it(*I sell carrots*). It will not match with the statement *'I sell vegetables'*

2.2 Improvement Strategy:

The system will be improved to match with the statements that has related terms, in addition to exact match statements by using the wordnet. The hyponyms, hypernyms, sister terms, meronyms and synsets will be employed to increase the efficiency of the system and thereby helps the consumer to get matched with the service provider who could most probably satisfy the need.

2.3 Examples:

Consumer Statements

I need a latte

I want to buy carrots

Service Provider Statements

I sell vegetables

We sell espresso

In the above examples

(i)Latte – one type of coffee where espresso is another type of coffee. These two gets matched by retrieving all the related forms of latte from wordnet.(espresso found as hypernym of latte in wordnet).

(ii)Vegetables are found as hypernym of Carrots in wordnet and is used to find this particular match.

2.4 Programming Tools:

Programming Language: Java

Packages used: Stanford POSTagger , Wordnet

2.5 Architecture Diagram:

