

Public Abstract

First Name:Akshay

Middle Name:

Last Name:Jain

Adviser's First Name:James

Adviser's Last Name:Keller

Co-Adviser's First Name:

Co-Adviser's Last Name:

Graduation Term:SP 2015

Department:Electrical Engineering

Degree:MS

Title:TEXTUAL SUMMARIZATION OF DATA USING LINGUISTIC PROTOFORM SUMMARIES

With the unseen quantities of data being generated in all walks of life, varying from social media to health domain, introduction of new techniques in order to better understand this information content is imperative. Contrary to the popular methods such as visualization or statistical summarization which requires the user to adapt to the technology, summarization of this information in the language pertaining to the audience, has recently gained a lot of traction. This thesis deals with textual summarization of data using Linguistic Protoform Summaries (LPS). We start by studying the existing techniques present in the literature to produce LPS, propose a new method and demonstrate its robustness with a mathematical proof. The usefulness of LPS is then illustrated with a novel application in the healthcare domain where the textual summaries are tailored to a clinical population. This is followed by an extensive study on the use of LPS as features in order to process data. There, we present our thoughts on the ways to handle LPS as data features and provide reasoning of this choice. We illustrate this with a real data example where we find a prototypical set of days of the activity of people living in an eldercare facility. Throughout this thesis we design various sets of experiments with synthetic data in order to explain the details of techniques presented.