ABSTRACT

Lease abstraction is the method of compartmentalization of key data from a lease document. Lease document for a property contains key business, money, and legal data about a property. A lease abstract report contains details concerning the property location and basic lease details, price schedules, key events, terms and conditions, automobile parking arrangements, and landowner and tenant obligations. Abstracting a true estate contract into electronic type facilitates easy access to key data, exchanging the tedious method of reading the whole contents of the contract every time. Language process may be used for data extraction and abstraction of knowledge from lease documents.
INTRODUCTION

Text Classification

Text classification is that the method of assignment tags or classes to text consistent with its content. It’s one amongst the elemental tasks in language process (NLP) with broad applications like sentiment analysis, topic labeling, spam detection, and intent detection (Esuli & Sebastiani, 2013). Unstructured data in the form of text is everywhere: emails, chats, web pages, social media, support tickets, survey responses, and more. Text can be an extremely rich source of information, but extracting insights from it can be hard and time-consuming due to its unstructured nature. Businesses are turning to text classification for structuring text in a fast and cost-efficient way to enhance decision-making and automate processes. With the help of text classification it’s able to classify between the documents whether it is lease documents for any office or for house.

Topic Modeling

Topic modeling could be a kind of statistical modeling for locating the abstract “topics” that occur in an exceedingly assortment of documents. Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) is associate degree example of topic model and is employed to classify text in an exceedingly document to a selected topic. It builds a subject per document model and words per topic model, sculptured as Dirichlet distributions. It is used to discover the topics that occur in the document

Information Extraction

Information Extraction (Hakkani-Tür, Ji, & Grishman, 2007) refers to the machine-controlled extraction of structured data like entities, relationships between entities, and attributes describing entities from unstructured sources. This enables a lot of richer styles of queries on the luxuriant unstructured sources than potential with keyword searches alone. When structured and unstructured data co-exist, information extraction makes it possible to integrate the two types of sources and pose queries spanning them.

The extraction of structure from creaking, unstructured sources may be a difficult task that has engaged a veritable community of researchers for over 20 years currently. With roots in the Natural Language Processing (NLP) community, the topic of structure extraction now engages many different communities spanning machine learning, information retrieval, database, web, and document analysis. Early extraction tasks were concentrated around the identification of named entities, like
Semantic Similarity Using Register Linear Question Classification (RLQC) for Question Classification
www.igi-global.com/chapter/semantic-similarity-using-register-linear-question-classification-rlqc-for-question-classification/245086?camid=4v1a

Areas of Narratives or Narrative Genres
(2020). Toward an Integrated Approach to Narrative Generation: Emerging Research and Opportunities (pp. 59-161).
www.igi-global.com/chapter/areas-of-narratives-or-narrative-genres/241120?camid=4v1a