Active Learning For Hierarchical Text Classi Cation

Yeah, reviewing a ebook active learning for hierarchical text classi cation could be credited with your close friends listings. This is just one of the solutions for you to be successful. As understood, expertise does not suggest that you have fabulous points.

Comprehending as without difficulty as pact even more than additional will have the funds for each success. bordering to, the pronouncement as without difficulty as perception of this active learning for hierarchical text classi cation can be taken as skillfully as picked to act.

Active Learning For Hierarchical Text
active learning in hierarchical text classification. Moreover, we explore how to utilize the hierarchical relation to further improve active learning. Accordingly, several leveraging strategies and heuristics are devised. According to our experiments, active learning under our framework significantly outperforms the

---

Active Learning for Hierarchical Text Classification
Active learning has been shown to reduce the training examples significantly, but it has not been applied to hierarchical text classification due to several technical challenges. In this paper, we study active learning for hierarchical text classification.

Active Learning for Hierarchical Text Classification ...
As far as we know, this is the first work that studies active learning in hierarchical text classification with promising results. Hierarchical text classification plays an important role in many real-world applications, such as webpage topic classification, product categorization and user feedback classification.

[PDF] Active Learning for Hierarchical Text Classification ...
Recently, a few works applies active learning methods to hierarchical classification. However, these methods fo-cus on the multi-class case, and regard only leaf nodes as labels in the hierarchal tree. Chenget al: [2012] embed the label hierarchy and training data into a latent semantic space, and propose a uncertainty strategy based on the se-

---

Cost-Effective Active Learning for Hierarchical Multi ...
Active Learning For Hierarchical Text Classi Cation Author: www.infraredtraining.com.br-2020-11-12T00:00:00+00:01 Subject: Active Learning For Hierarchical Text Classi Cation Keywords: active, learning, for, hierarchical, text, classi, cation Created Date: 11/12/2020 2:36:23 PM

Active Learning For Hierarchical Text Classi Cation
Download PDF: Sorry, we are unable to provide the full text but you may find it at the following location(s): http://cling.csd.uwo.ca/papers... (external link)

Active Learning for Hierarchical Text Classification - CORE

Active learning (AL) is a subfield of machine learning which addresses methods to build models with fewer, but more representative instances.

(PDF) Active learning for hierarchical multi-label ...
Active Learning, Hierarchical Classiflcation, Label Tree Embedding 1. INTRODUCTION Obtaining labels is an expensive or time-consuming pro-cess, especially for large scale multi-class classiflcation prob-lems. Active learning is proposed to make the learning task more e–cient [12], by intelligently choosing speciflc unla-

---

On active learning in hierarchical classification
active learning is helpful. The first has to do with effi-cient search through the hypothesis space. Each time a new label is seen, the set of plausible classifiers (those roughly consistent with the labels seen so far) shrinks somewhat. Using active learning, one can explicitly select points whose labels will shrink this set as fast as possible.

Hierarchical Sampling for Active Learning
A digital learning space for your pupils and a toolkit for you, so that you can search, plan, allocate and assess all in one place.

ActiveLearn: Login
Active learning (AL) is a subfield of machine learning which addresses methods to build models with fewer, but more representative instances. Even though AL has been vastly

studied, it has not been thoroughly investigated in hierarchical multi-label classification, a learning task where multiple class labels can be assigned to an instance and these labels are hierarchically structured.

Active learning for hierarchical multi-label ...
Where To Download Active Learning For Hierarchical Text Classi Cationdownload paywalled content for free including PDF downloads for the stuff on Elsevier ' s Science Direct website. Even though the site continues to face legal issues due to the pirated access provided to books and articles, the site is still functional through various domains.

Active Learning For Hierarchical Text Classi Cation
Active learning has been shown to reduce the training examples significantly, but it has not been applied to hierarchical text classification due to several technical challenges. In this paper, we study active learning for hierarchical text classification.

CiteSeerX — Active Learning for Hierarchical Text ...
This paper presents a new approach to hierarchical document classification that we call Hierarchical Deep Learning for Text classification (HDLTex). 1 1 1 HDLTex is shared as an open source tool at https://github.com/kk7nc/HDLTex HDLTex combines deep learning architectures to allow both overall and specialized learning by level of the document hierarchy. This paper reports our experiments with HDLTex, which exhibits improved accuracy over traditional document classification

methods.

HDLTex: Hierarchical Deep Learning for Text Classification An Active Learning Framework for Hyperspectral Image Classification Using Hierarchical Segmentation. Abstract: Augmenting spectral data with spatial information for image classification has recently gained significant attention, as classification accuracy can often be improved by extracting spatial information from neighboring pixels. In this paper, we propose a new framework in which active learning (AL) and hierarchical segmentation (HSeg) are combined for spectral-spatial classification ...

An Active Learning Framework for Hyperspectral Image ... We are pleased to announce the 4th edition of the Large Scale Hierarchical Text Classification (LSHTC) Challenge. The LSHTC Challenge is a hierarchical text classification competition, using very large datasets. Hierarchies are becoming ever more popular for the organization of text documents, particularly on the Web.

Large Scale Hierarchical Text Classification ¦ Kaggle We investigate active learning by pairwise similarity over the leaves of trees originating from hierarchical clustering procedures. In the realizable setting, we provide a full characterization of the number of queries needed to achieve perfect reconstruction of the tree cut.

Flattening a Hierarchical Clustering through Active Learning

Abstract Most of the existing active learning algorithms assume all the category labels as independent or consider them in a "flat" structure. However, in reality, there are many applications in which the set of possible labels are often organized in a hierarchical structure.

On active learning in hierarchical classification ...
Request PDF ¦ On active learning in hierarchical classification ¦ Most of the existing active learning algorithms assume all the category labels as independent or consider them in a "flat" structure.

The two-volume set LNAI 7301 and 7302 constitutes the refereed proceedings of the 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2012, held in Kuala Lumpur, Malaysia, in May 2012. The total of 20 revised full papers and 66 revised short papers were carefully reviewed and selected from 241 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD-related areas. The papers are organized in topical sections on supervised learning: active, ensemble, rare-class and online; unsupervised learning: clustering, probabilistic modeling in the first volume and on pattern mining: networks, graphs, time-series and outlier detection, and data manipulation: pre-processing and dimension reduction in the second volume.

The two-volume set LNAI 7818 + LNAI 7819 constitutes the refereed proceedings of the 17th Pacific-Asia Conference on

Knowledge Discovery and Data Mining, PAKDD 2013, held in Gold Coast, Australia, in April 2013. The total of 98 papers presented in these proceedings was carefully reviewed and selected from 363 submissions. They cover the general fields of data mining and KDD extensively, including pattern mining, classification, graph mining, applications, machine learning, feature selection and dimensionality reduction, multiple information sources mining, social networks, clustering, text mining, text classification, imbalanced data, privacy-preserving data mining, recommendation, multimedia data mining, stream data mining, data preprocessing and representation.

Comprehensive Coverage of the Entire Area of Classification Research on the problem of classification tends to be fragmented across such areas as pattern recognition, database, data mining, and machine learning. Addressing the work of these different communities in a unified way, Data Classification: Algorithms and Applications explores the underlying algorithms of classification as well as applications of classification in a variety of problem domains, including text, multimedia, social network, and biological data. This comprehensive book focuses on three primary aspects of data classification: Methods-The book first describes common techniques used for classification, including probabilistic methods, decision trees, rule-based methods, instance-based methods, support vector machine methods, and neural networks. Domains-The book then examines specific methods used for data domains such as multimedia, text, time-series, network, discrete sequence, and uncertain data. It also covers large data sets and data streams due to the recent importance of the big data paradigm. Variations-The book concludes with insight on variations of the classification process. It discusses

ensembles, rare-class learning, distance function learning, active learning, visual learning, transfer learning, and semi-supervised learning as well as evaluation aspects of classifiers.

The two-volume set LNAI 7301 and 7302 constitutes the refereed proceedings of the 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2012, held in Kuala Lumpur, Malaysia, in May 2012. The total of 20 revised full papers and 66 revised short papers were carefully reviewed and selected from 241 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD-related areas. The papers are organized in topical sections on supervised learning: active, ensemble, rare-class and online; unsupervised learning: clustering, probabilistic modeling in the first volume and on pattern mining: networks, graphs, time-series and outlier detection, and data manipulation: pre-processing and dimension reduction in the second volume.

Over the last two decades, researchers are looking at imbalanced data learning as a prominent research area. Many critical real-world application areas like finance, health, network, news, online advertisement, social network media, and weather have imbalanced data, which emphasizes the research necessity for real-time implications of precise fraud/defaulter detection, rare disease/reaction prediction, network intrusion detection, fake news detection, fraud advertisement detection, cyber bullying identification, disaster events prediction, and more. Machine learning algorithms are based on the heuristic of equally-distributed balanced data and provide the biased result towards the majority data class, which is not

acceptable considering imbalanced data is omnipresent in real-life scenarios and is forcing us to learn from imbalanced data for foolproof application design. Imbalanced data is multifaceted and demands a new perception using the novelty at sampling approach of data preprocessing, an active learning approach, and a cost perceptive approach to resolve data imbalance. Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance offers new aspects for imbalanced data learning by providing the advancements of the traditional methods, with respect to big data, through case studies and research from experts in academia, engineering, and industry. The chapters provide theoretical frameworks and the latest empirical research findings that help to improve the understanding of the impact of imbalanced data and its resolving techniques based on data preprocessing, active learning, and cost perceptive approaches. This book is ideal for data scientists, data analysts, engineers, practitioners, researchers, academicians, and students looking for more information on imbalanced data characteristics and solutions using varied approaches.

The key idea behind active learning is that a machine learning algorithm can perform better with less training if it is allowed to choose the data from which it learns. An active learner may pose "queries," usually in the form of unlabeled data instances to be labeled by an "oracle" (e.g., a human annotator) that already understands the nature of the problem. This sort of approach is well-motivated in many modern machine learning and data mining applications, where unlabeled data may be abundant or easy to come by, but training labels are difficult, time-consuming, or expensive to obtain. This book is a general introduction to active learning. It outlines several scenarios in which queries

might be formulated, and details many query selection algorithms which have been organized into four broad categories, or "query selection frameworks." We also touch on some of the theoretical foundations of active learning, and conclude with an overview of the strengths and weaknesses of these approaches in practice, including a summary of ongoing work to address these open challenges and opportunities. Table of Contents: Automating Inquiry / Uncertainty Sampling / Searching Through the Hypothesis Space / Minimizing Expected Error and Variance / Exploiting Structure in Data / Theory / Practical Considerations

Web search engines have been adopted by most universities fo r searching webpages in their own domains. Basically, a user sends keywords to the search e ngine and the search engine returns a flat ranked list of webpages. However, in universit y search, user queries are usually related to topics. Simple keyword queries are often insu ffi cient to express topics as keywords. On the other hand, most E-commerce sites allow users to brows e and search products in various hierarchies. It would be ideal if hierarchical browsing and keyword search can be seamlessly combined for university search engines. The main di ffi culty is to automatically classify and rank a massive number of webpages into the topic hierarchies for universities. In this thesis, we use machine learning and data mining techn iques to build a novel hybrid search engine with integrated hierarchies for universitie s, called SEEU (S earch E ngine with hi E rarchy for U niversities). Firstly, we study the problem of e ff ective hierarchical webpage classification. We develop a parallel webpage classification system based on Support Ve ctor Machines. With extensive experiments on the well-known ODP (Open Directory Project) dataset, we empirically demon- strate that our hierarchical classification

system is very e ff ective and outperforms the traditional flat classification approaches significantly. Secondly, we study the problem of integrating hierarchical classification into the ranking system of keywords-based search engines. We propose a novel ranking framework, called ERIC (E nhanced R anking by h I erarchical C lassification), for search engines with hierarchies. Experimental results on four large-scale TREC (Text REtrie val Conference) web search datasets show that our ranking system with hierarchical classificati on outperforms the traditional flat keywords-based search methods significantly. Thirdly, we propose a novel active learning framework to imp rove the performance of hierarchical classification, which is important for ranking w ebpages in hierarchies. From our experiments on the benchmark text datasets, we find that our a ctive learning framework can achieve good classification performance yet save a consider able number of labeling e ff ort com- pared with the state-of-the-art active learning methods fo r hierarchical text classification. Fourthly, based on the proposed classification and ranking m ethods, we present a novel hierarchical classification framework for mining academic topics from university webpages. We build an academic topic hierarchy based on the commonly ac cepted Wikipedia academic disciplines. Based on this hierarchy, we train a hierarchic al classifier and apply it to mine academic topics. According to our comprehensive analysis, the academic topics mined by our method are reasonable and consistent with the real-world to pic distribution in universities. Finally, we combine all the proposed techniques together an d implement the SEEU search engine. According to two usability studies conducted in the ECE and the CS departments at our university, SEEU is favored by the majority of participa nts. To conclude, the main contribution of this thesis is a novel s earch engine,

called SEEU, for universities. We discuss the challenges toward building SE EU and propose e ff ective machine learning and data mining methods to tackle them. With extens ive experiments on well-known benchmark datasets and real-world university webpage data sets, we demonstrate that our sys- tem is very e ff ective. In addition, two usability studies of SEEU in our uni versity show that SEEU has a great promise for university search.

The key idea behind active learning is that a machine learning algorithm can perform better with less training if it is allowed to choose the data from which it learns. An active learner may pose "queries," usually in the form of unlabeled data instances to be labeled by an "oracle" (e.g., a human annotator) that already understands the nature of the problem. This sort of approach is well-motivated in many modern machine learning and data mining applications, where unlabeled data may be abundant or easy to come by, but training labels are difficult, time-consuming, or expensive to obtain. This book is a general introduction to active learning. It outlines several scenarios in which queries might be formulated, and details many query selection algorithms which have been organized into four broad categories, or "query selection frameworks." We also touch on some of the theoretical foundations of active learning, and conclude with an overview of the strengths and weaknesses of these approaches in practice, including a summary of ongoing work to address these open challenges and opportunities. Table of Contents: Automating Inquiry / Uncertainty Sampling / Searching Through the Hypothesis Space / Minimizing Expected Error and Variance / Exploiting Structure in Data / Theory / Practical Considerations

The 3-volume set LNAI 12712-12714 constitutes the

proceedings of the 25th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, PAKDD 2021, which was held during May 11-14, 2021. The 157 papers included in the proceedings were carefully reviewed and selected from a total of 628 submissions. They were organized in topical sections as follows: Part I: Applications of knowledge discovery and data mining of specialized data; Part II: Classical data mining; data mining theory and principles; recommender systems; and text analytics; Part III: Representation learning and embedding, and learning from data.