Automatic Feature Selection for Named Entity Recognition Using Genetic Algorithm

Huong Thanh Le
Hanoi University of Science and Technology
Hanoi, Vietnam
huongle@soict.hut.edu.vn

Luan Van Tran
Hanoi University of Science and Technology
Hanoi, Vietnam
tranvanluan7@gmail.com

ABSTRACT

This paper presents a feature selection approach for named entity recognition using genetic algorithm. Different aspects of genetic algorithm including computational time and criteria for evaluating an individual (i.e., size of the feature subset and the classifier's accuracy) are analyzed in order to optimize its learning process. Two machine learning algorithms, k-Nearest Neighbor and Conditional Random Fields, are used to calculate the accuracy of the named entity recognition system. To evaluate the effectiveness of our genetic algorithm, feature subsets returning by our proposed genetic algorithm are compared to feature subsets returning by a hill climbing algorithm and a backward one. Experimental results show that feature subsets obtained by our genetic algorithm is much smaller than the original feature set without losing of predictive accuracy. Furthermore, these feature subsets result in higher classifier's accuracies than that of the hill climbing algorithm and the backward one.

Categories and Subject Descriptors

G.2.1 [Combinatorics] - *Combinatorial algorithms, Counting problems*

General Terms

Algorithms, Measurement, Performance, Design, Reliability, Experimentation, Languages, Theory.

Keywords

Optimizing, genetic algorithm, feature selection, named entity recognition.

1. INTRODUCTION

Named Entity Recognition (NER) is the task of locating and classifying atomic elements in text into predefined categories (e.g., people, organizations, locations, expressions of time, quantities, etc.). NER is useful in many applications such as search engine, question-answering systems, etc.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Conference'10, Month 1–2, 2010, City, State, Country.
Copyright 2010 ACM 1-58113-000-0/00/0010…$10.00.

Most NER's systems concentrate on machine learning approaches, whose performance relies on features being used. Selecting a subset of good features is an optimization problem of multiple objectives, such as number of features and accuracy. Using a small feature subset requires less computational time than using a larger one. However, the system's accuracy is lower than using a larger, suitable set of features. Another problem is, using more features increases the system complexity, but does not always increase the system's accuracy. A bad feature may even degrade the performance of the system. Most machine learning approaches for NER determine the features manually. In this paper, we propose a genetic algorithm (GA) to automatically select a minimal subset of features that can produce high accuracy for the NER system. Different aspects are discussed in order to optimize the GA learning process.

The rest of this paper is organized as follows. Section 2 introduces related work of feature selection using GA and the contribution of this paper. Section 3 describes our approach to apply GA in optimizing the set of features for NER system. Experimental results are discussed in Section 4. Finally, Section 5 concludes the paper and indicates some future research directions.

2. RELATED WORK

Feature selection is the process of selecting an optimal subset from original features. In a NER system, an optimal subset of features is the smallest subset that most contribute towards accuracy. The smaller size of the feature subset is, the faster the NER system runs. However, the size of feature's subsets is not proportional with the accuracy. Therefore, this is an optimization problem of multiple objectives.

Carrying out an exhaustive search in the feature space has a high computational cost, since the total search space for M features is 2^M, which is a NP-hard problem. Therefore, heuristic search is often used to deal with this problem.

Two approaches often used in single solution are forward selection [15,16] and backward selection [1,8]. The forward selection starts with no feature in the model. It tests the addition of each feature using a chosen model comparison criterion and adding the feature (if any) that improves the model the most. This process repeats until none improves the model. The backward selection starts with all candidate variables. It tests the deletion of each variable using a chosen model comparison criterion and deleting the variable (if any) that improves the model the most by being deleted. This process repeats until no further improvement is possible. The disadvantage of these strategies is, the searching process cannot go back to previous features. Therefore, if a

Automatic Feature Selection For Named Entity Recognition

Mohd Saberi Mohamad,Miguel P. Rocha,Florentino Fdez-Riverola,Francisco J. Domínguez Mayo,Juan F. De Paz

Automatic Feature Selection For Named Entity Recognition:

Immerse yourself in heartwarming tales of love and emotion with Explore Love with is touching creation, Tender Moments: **Automatic Feature Selection For Named Entity Recognition** . This emotionally charged ebook, available for download in a PDF format (PDF Size: *), is a celebration of love in all its forms. Download now and let the warmth of these stories envelop your heart.

https://py.bijouxmedusa.com/book/virtual-library/Documents/92_1709_healthy_recipes_case_study_for_creators_92_1011_healthy_recipes.pdf

Table of Contents Automatic Feature Selection For Named Entity Recognition

1. Understanding the eBook Automatic Feature Selection For Named Entity Recognition
    - The Rise of Digital Reading Automatic Feature Selection For Named Entity Recognition
    - Advantages of eBooks Over Traditional Books
2. Identifying Automatic Feature Selection For Named Entity Recognition
    - Exploring Different Genres
    - Considering Fiction vs. Non-Fiction
    - Determining Your Reading Goals
3. Choosing the Right eBook Platform
    - Popular eBook Platforms
    - Features to Look for in an Automatic Feature Selection For Named Entity Recognition
    - User-Friendly Interface
4. Exploring eBook Recommendations from Automatic Feature Selection For Named Entity Recognition
    - Personalized Recommendations
    - Automatic Feature Selection For Named Entity Recognition User Reviews and Ratings
    - Automatic Feature Selection For Named Entity Recognition and Bestseller Lists
5. Accessing Automatic Feature Selection For Named Entity Recognition Free and Paid eBooks
    - Automatic Feature Selection For Named Entity Recognition Public Domain eBooks
    - Automatic Feature Selection For Named Entity Recognition eBook Subscription Services

- Automatic Feature Selection For Named Entity Recognition Budget-Friendly Options

6. Navigating Automatic Feature Selection For Named Entity Recognition eBook Formats
   - ePub, PDF, MOBI, and More
   - Automatic Feature Selection For Named Entity Recognition Compatibility with Devices
   - Automatic Feature Selection For Named Entity Recognition Enhanced eBook Features
7. Enhancing Your Reading Experience
   - Adjustable Fonts and Text Sizes of Automatic Feature Selection For Named Entity Recognition
   - Highlighting and Note-Taking Automatic Feature Selection For Named Entity Recognition
   - Interactive Elements Automatic Feature Selection For Named Entity Recognition
8. Staying Engaged with Automatic Feature Selection For Named Entity Recognition
   - Joining Online Reading Communities
   - Participating in Virtual Book Clubs
   - Following Authors and Publishers Automatic Feature Selection For Named Entity Recognition
9. Balancing eBooks and Physical Books Automatic Feature Selection For Named Entity Recognition
   - Benefits of a Digital Library
   - Creating a Diverse Reading Collection Automatic Feature Selection For Named Entity Recognition
10. Overcoming Reading Challenges
    - Dealing with Digital Eye Strain
    - Minimizing Distractions
    - Managing Screen Time
11. Cultivating a Reading Routine Automatic Feature Selection For Named Entity Recognition
    - Setting Reading Goals Automatic Feature Selection For Named Entity Recognition
    - Carving Out Dedicated Reading Time
12. Sourcing Reliable Information of Automatic Feature Selection For Named Entity Recognition
    - Fact-Checking eBook Content of Automatic Feature Selection For Named Entity Recognition
    - Distinguishing Credible Sources
13. Promoting Lifelong Learning
    - Utilizing eBooks for Skill Development
    - Exploring Educational eBooks
14. Embracing eBook Trends

- Integration of Multimedia Elements
- Interactive and Gamified eBooks

Automatic Feature Selection For Named Entity Recognition Introduction

In the digital age, access to information has become easier than ever before. The ability to download Automatic Feature Selection For Named Entity Recognition has revolutionized the way we consume written content. Whether you are a student looking for course material, an avid reader searching for your next favorite book, or a professional seeking research papers, the option to download Automatic Feature Selection For Named Entity Recognition has opened up a world of possibilities. Downloading Automatic Feature Selection For Named Entity Recognition provides numerous advantages over physical copies of books and documents. Firstly, it is incredibly convenient. Gone are the days of carrying around heavy textbooks or bulky folders filled with papers. With the click of a button, you can gain immediate access to valuable resources on any device. This convenience allows for efficient studying, researching, and reading on the go. Moreover, the cost-effective nature of downloading Automatic Feature Selection For Named Entity Recognition has democratized knowledge. Traditional books and academic journals can be expensive, making it difficult for individuals with limited financial resources to access information. By offering free PDF downloads, publishers and authors are enabling a wider audience to benefit from their work. This inclusivity promotes equal opportunities for learning and personal growth. There are numerous websites and platforms where individuals can download Automatic Feature Selection For Named Entity Recognition. These websites range from academic databases offering research papers and journals to online libraries with an expansive collection of books from various genres. Many authors and publishers also upload their work to specific websites, granting readers access to their content without any charge. These platforms not only provide access to existing literature but also serve as an excellent platform for undiscovered authors to share their work with the world. However, it is essential to be cautious while downloading Automatic Feature Selection For Named Entity Recognition. Some websites may offer pirated or illegally obtained copies of copyrighted material. Engaging in such activities not only violates copyright laws but also undermines the efforts of authors, publishers, and researchers. To ensure ethical downloading, it is advisable to utilize reputable websites that prioritize the legal distribution of content. When downloading Automatic Feature Selection For Named Entity Recognition, users should also consider the potential security risks associated with online platforms. Malicious actors may exploit vulnerabilities in unprotected websites to distribute malware or steal personal information. To protect themselves, individuals should ensure their devices have reliable antivirus software installed and validate the legitimacy of the websites they are downloading from. In conclusion, the ability to download Automatic Feature Selection For Named Entity Recognition has transformed the way we access information. With the convenience, cost-effectiveness, and accessibility it offers, free PDF

downloads have become a popular choice for students, researchers, and book lovers worldwide. However, it is crucial to engage in ethical downloading practices and prioritize personal security when utilizing online platforms. By doing so, individuals can make the most of the vast array of free PDF resources available and embark on a journey of continuous learning and intellectual growth.

FAQs About Automatic Feature Selection For Named Entity Recognition Books

1. Where can I buy Automatic Feature Selection For Named Entity Recognition books? Bookstores: Physical bookstores like Barnes & Noble, Waterstones, and independent local stores. Online Retailers: Amazon, Book Depository, and various online bookstores offer a wide range of books in physical and digital formats.
2. What are the different book formats available? Hardcover: Sturdy and durable, usually more expensive. Paperback: Cheaper, lighter, and more portable than hardcovers. E-books: Digital books available for e-readers like Kindle or software like Apple Books, Kindle, and Google Play Books.
3. How do I choose a Automatic Feature Selection For Named Entity Recognition book to read? Genres: Consider the genre you enjoy (fiction, non-fiction, mystery, sci-fi, etc.). Recommendations: Ask friends, join book clubs, or explore online reviews and recommendations. Author: If you like a particular author, you might enjoy more of their work.
4. How do I take care of Automatic Feature Selection For Named Entity Recognition books? Storage: Keep them away from direct sunlight and in a dry environment. Handling: Avoid folding pages, use bookmarks, and handle them with clean hands. Cleaning: Gently dust the covers and pages occasionally.
5. Can I borrow books without buying them? Public Libraries: Local libraries offer a wide range of books for borrowing. Book Swaps: Community book exchanges or online platforms where people exchange books.
6. How can I track my reading progress or manage my book collection? Book Tracking Apps: Goodreads, LibraryThing, and Book Catalogue are popular apps for tracking your reading progress and managing book collections. Spreadsheets: You can create your own spreadsheet to track books read, ratings, and other details.
7. What are Automatic Feature Selection For Named Entity Recognition audiobooks, and where can I find them? Audiobooks: Audio recordings of books, perfect for listening while commuting or multitasking. Platforms: Audible, LibriVox, and Google Play Books offer a wide selection of audiobooks.
8. How do I support authors or the book industry? Buy Books: Purchase books from authors or independent bookstores. Reviews: Leave reviews on platforms like Goodreads or Amazon. Promotion: Share your favorite books on social media

or recommend them to friends.

9. Are there book clubs or reading communities I can join? Local Clubs: Check for local book clubs in libraries or community centers. Online Communities: Platforms like Goodreads have virtual book clubs and discussion groups.
10. Can I read Automatic Feature Selection For Named Entity Recognition books for free? Public Domain Books: Many classic books are available for free as theyre in the public domain. Free E-books: Some websites offer free e-books legally, like Project Gutenberg or Open Library.

Find Automatic Feature Selection For Named Entity Recognition :

92-1709 healthy recipes case study for creators 92-1011 healthy recipes
learning basics examples America 92-197 machine learning basics
92-1186 NFT marketplace guide USA 92-2836 NFT marketplace guide for
~~for entrepreneurs 92-1044 machine learning basics guide for small~~
92-726 startup funding guide for entrepreneurs 92-367 startup funding
92-569 TikTok marketing tutorial USA 92-1633 TikTok marketing tutorial
by step for startups 92-1761 blog monetization step by step for startups
beginners America 92-2900 freelancing online for beginners for
~~tools review USA 92-903 AI tools review United States 92-1591 AI tools~~
~~writing explained for startups 92-2126 resume writing for beginners USA~~
creators 92-2095 productivity hacks ideas for startups 92-381
growth step by step for startups 92-1371 Instagram growth strategies
startups 92-818 travel tips examples for creators 92-1726 travel tips
92-2526 blog monetization step by step for startups 92-436 blog
marketing case study for entrepreneurs 92-2743 TikTok marketing case

Automatic Feature Selection For Named Entity Recognition :

Strategic Planning For Success: Aligning People ... - Wiley Strategic Planning For Success: Aligning People ... - Wiley Strategic Planning For Success: Aligning... by Roger ... Useful, pragmatic, and proven tools and concepts, including needs assessment, needs analysis, and costs–consequences analysis. Strategic Planning for Success ... Strategic Planning For Success: Aligning People ... Strategic Planning for Success will show you how to define, deliver, develop, and promote

genuine performance improvement within your organization. --This text ... Strategic planning for success; aligning people TITLE: Strategic planning for success; aligning people, performance, and payoffs. AUTHOR: Kaufman, Roger et al. PUBLISHER: Jossey-Bass ... Strategic Planning for Success Welcome to Strategic Planning for Success: Aligning People, Performance, and Payoffs. This is a practical and pragmatic book with cases-in-point, guides, job. Strategic Planning For Success: Aligning People, ... Strategic Planning for Success offers you a pragmatic guide to the design and development of practical and pragmatic strategic thinking and organizational ... Strategic Planning For Success: Aligning People, Performance ... Strategic Planning for Success offers you a pragmatic guide to the design and development of practical and pragmatic strategic thinking and organizational ... Book Review: Strategic Planning for Success: Aligning ... Roger Kaufman, Hugh Oakley-Browne, Ryan Watkins, and Doug Leigh As I read this book, my first reaction was, although it covered a lot of territory with ... Strategic planning for success - Vanderbilt Libraries Catalog Strategic planning for success : aligning people, performance, and payoffs / Roger Kaufman . ... Strategic planning for success : aligning people, performance ... Strategic Planning for Success: Aligning People ... Mar 6, 2003 — Strategic Planning for Success offers you a pragmatic guide to the design and development of practical and pragmatic strategic thinking and ... COMP XM Flashcards Study with Quizlet and memorize flashcards containing terms like Segment/Perf/Size, Prices between each round, Price for each product and more. COMP XM Exam : r/Capsim The questions are a bit hard and change a lot from exam to exam so do not trust too much the keys you find online, most of them are about ... Board Query 1 Questions and Answers for FINAL COMP ... Aug 4, 2023 — Board Query 1 Questions and Answers for FINAL COMP XM EXAM. CompXM Capsim Examination Notes - BOD QUIZ Q1) ... Q1) Rank the following companies from high to low cumulative profit, (in descending order, 1=highest,. 4=lowest). Answer 1) From Selected Financial Statistic ... Board Query 1 Questions for FINAL COMP XM EXAM.pdf The rise in the labour cost increase the price of the Jacket and the quality of the supply remain unchanged. Is this a violation of the law of supply? Explain. COMPXM answers 2024 This article provides COMPXM answers 2024 template. It offers answers for round 1 and guide make decisions for remaining comp XM rounds. This comp-xm guide ... 7 Comp-XM The Comp-XM Competency Exam is built around a simulation similar to Capstone and Foundation. ... This makes the questions comparable but the answers unique. Health Care USA: Understanding Its... by Sultz, Harry Book details ; ISBN-10. 1284002802 ; ISBN-13. 978-1284029888 ; Edition. 8th ; Publisher. Jones & Bartlett Learning ; Publication date. July 19, 2013. Health Care USA: Understanding Its Organization and ... Health Care USA, Eighth Edition Includes Navigate Advantage Access, offers students of health administration, public health, medicine, and related fields a ... Health Care USA: Understanding Its Organization and ... Health Care USA: Understanding Its Organization and Delivery, 8th Edition by Sultz, Harry - ISBN 10: 1284029883 - ISBN 13: 9781284029888 - Jones & Bartlett ... Health Care USA: Understanding Its Organization and ... Health Care USA, Eighth Edition Includes Navigate Advantage Access, offers students of health administration, public health,

medicine, and related fields a ... Health Care USA 8th edition 9781284029888 1284029883 Health Care USA: Understanding Its Organization and Delivery · 8th edition · 978-1284029888 · Paperback/softback · Jones & Bartlett (7/19/2013). Health Care USA: Understanding Its Organization and ... Health Care USA, Eighth Edition, offers students of health administration, public health, medicine, and related fields a wide-ranging overview of America's ... Sultz and Young's Health Care USA: Understanding Its ... Sultz and Young's Health Care USA: Understanding Its Organization and Deliveryselected product title. Tenth Edition. James A. Johnson, PhD, MPA, MSc; Kimberly ... Health Care USA: Understanding Its Organization and ... Health Care USA: Understanding Its Organization and Delivery, 8th Edition ; No reviews yet Write a review ; Subscribe to Discover Books. Exclusive discount codes, ... Health Care USA book by Kristina M Young Health Care USA: Understanding Its Organization and Delivery, 8th Edition. Kristina M. Young, Harry A. Sultz. Health Care USA: Understanding Its Organization and ... Health Care USA: Understanding Its Organization and Delivery, 8th Edition by Su ; Condition. Brand New ; Quantity. 1 available ; Item Number. 335124557461 ; ISBN.