

1. Record Nr.	UNINA9910841863203321
Autore	Tan Ying
Titolo	Data Mining and Big Data [[electronic resource] ] : 8th International Conference, DMBD 2023, Sanya, China, December 9–12, 2023, Proceedings, Part I // edited by Ying Tan, Yuhui Shi
Pubbl/distr/stampa	Singapore : , : Springer Nature Singapore : , : Imprint : Springer, , 2024
ISBN	981-9708-37-0
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (297 pages)
Collana	Communications in Computer and Information Science, , 1865-0937 ; ; 2017
Altri autori (Persone)	ShiYuhui
Disciplina	005.3
Soggetti	Application software Artificial intelligence Image processing - Digital techniques Computer vision Computer and Information Systems Applications Artificial Intelligence Computer Imaging, Vision, Pattern Recognition and Graphics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Organization -- Contents - Part I -- Contents - Part II -- Data Mining Methods -- Data Analytics Methods in Human Resource Demand Forecasting -- 1 Introduction -- 2 Related work -- 2.1 Human Resource Demand Forecasting methods -- 2.2 Multiple Regression Analysis -- 2.3 Data Mining methods -- 3 Data Analytics Methods in Human Resource Demand Forecasting -- 3.1 Multiple Regression Analysis Method -- 3.2 BP Neural Network Algorithm -- 4 Experimental Analysis -- 4.1 Result Analysis of Multiple Linear Regression Forecasting for Human Resources -- 4.2 Result Analysis of BP Neural Network Forecasting for Human Resources -- References -- A Localization Correction Algorithm of Location-Based Services Based on Point Clustering -- 1 Introduction -- 2 Cell Phone Location and Location Service Bias -- 2.1 Cell Phone Positioning -- 2.2 Location Service Bias -- 3 Core Technology -- 3.1 Clustering Algorithm -- 3.2 Implementation of Localization in ANDROID Systems -- 3.3 Expectation

Model for Location Data -- 4 Positioning Data Acquisition Program Design and Implementation -- 4.1 Positional Data Structure -- 4.2 Positioning Process -- 5 Experimental Analysis -- 6 Summarize -- References -- Comparison of Prediction Methods on Large-Scale and Long-Term Online Live Streaming Data -- 1 Introduction -- 2 Dataset and Methods -- 2.1 Live Streaming Data -- 2.2 Prediction Methods -- 2.3 Experimental Design -- 3 Results -- 3.1 Data Overview and Temporal Analysis -- 3.2 Univariate Prediction -- 3.3 Feature Importance -- 4 Conclusions and Discussion -- References -- Forecasting Chinese Overnight Stock Index Movement Using Large Language Models with Market Summary -- 1 Introduction -- 2 Data -- 2.1 Market Summary -- 2.2 Overnight Stock Market Index Movement -- 3 Models -- 3.1 BERT -- 3.2 FinBERT -- 3.3 RoBERTa -- 3.4 LERT -- 3.5 MacBERT -- 3.6 PERT. -- 3.7 BART -- 3.8 Longformer -- 3.9 BigBird -- 3.10 GPT -- 3.11 Fine-tuning -- 3.12 Prompt of GPT -- 3.13 Evaluation Metrics -- 4 Empirical Results -- 5 Conclusions -- References -- A Practical Byzantine Fault Tolerance Algorithms Based on Randomized Mean Clustering, Trust and Credibility -- 1 Introduction -- 2 Related Works -- 3 CTPBFT Algorithm Design and Implementation -- 3.1 Node Trust Assessment -- 3.2 CTPBFT Divided Cluster Strategy -- 3.3 Node Credibility and Reward Curves -- 3.4 Improved Consensus Process -- 4 Reward Functions -- 4.1 Feasibility Analysis of the Reward Function -- 5 Experiment -- 5.1 Algorithm Complexity Analysis -- 5.2 Communication Overhead -- 5.3 Message Throughput -- 5.4 Transaction Delay -- 6 Conclusion -- References -- Improved Joint Distribution Adaptation for Fault Diagnosis -- 1 Introduction -- 2 Transfer Learning Problems -- 3 Proposed Method -- 3.1 Reconstruction with Weighted Source Domain -- 3.2 Modeling and Solving -- 4 Experiment Results and Comparison -- 4.1 Dataset -- 4.2 Results of Data Experiments -- 5 Conclusion -- References -- Pretrained Language Models and Applications -- A Unified Recombination and Adversarial Framework for Machine Reading Comprehension -- 1 Introduction -- 2 Related Work -- 2.1 Attention Mechanism Based Methods -- 2.2 Pre-trained Language Model Based Methods -- 3 Methods -- 3.1 Recombination Layer -- 3.2 Encoding Layer -- 3.3 Fusion Layer -- 3.4 Adversarial Learning -- 3.5 Model Training and Prediction -- 4 Experiment -- 4.1 Experimental Setup -- 4.2 Overall Results -- 4.3 Multi-type Problem Analysis -- 4.4 Analysis of Different Numbers of Candidates -- 5 Conclusion -- References -- Research on Data Mining Methods in the Field of Quality Problem Analysis Based on BERT Model -- 1 Introduction -- 2 Related Work -- 2.1 Data Related to Quality Problem Analysis -- 2.2 BERT Model -- 3 Method. -- 3.1 Pre-processing of Data -- 3.2 Named Entity Recognition Task -- 3.3 Relationship Extraction Task -- 4 Experiments -- 4.1 Construction of Dataset -- 4.2 Named Entity Recognition Task -- 4.3 Relationship Extraction Task -- 5 Conclusion -- References -- Cross-Language Text Search Algorithm Based on Context-Compatible Algorithms -- 1 Introduction -- 2 Research Status -- 2.1 Overview of the Linguistic Contextual Representation Model -- 2.2 Context-Compatible Algorithm -- 2.3 Participle Algorithm -- 3 Core Algorithm Implementation -- 3.1 Segmentation Algorithm Implementation -- 3.2 Word Co-sentence Algorithm Implementation -- 3.3 Context-Based Cross-Context Search Algorithm Implementation -- 4 Experiments -- 4.1 Experimentation with Segmentation Algorithms -- 4.2 Corpus Realization -- 4.3 Implementation of Cross-Language Search Algorithms -- 5 Summarize -- References -- PEKD: Joint Prompt-Tuning and Ensemble Knowledge Distillation Framework for Causal

Event Detection from Biomedical Literature -- 1 Introduction -- 2 Related Work -- 2.1 Event Relation Extraction Approach Based on Neural Network -- 2.2 Few-Shot Learning Methods for Neural Network Models -- 2.3 Compression Methods for Neural Network Models -- 3 Main Method -- 3.1 Problem Definition -- 3.2 Framework -- 4 Datasets and Experiments -- 4.1 Datasets -- 4.2 Evaluation Metric -- 4.3 Hyper-Parameters -- 4.4 Overall Performance -- 4.5 Ablation Study -- 5 Conclusion -- References -- Empirical Analysis on the Effectiveness of Pre-trained Models in the Identification of Physical Violence Against Women in Videos for a Multi-class Approach -- 1 Introduction -- 2 Theoretical Foundations -- 2.1 VGG16 -- 2.2 ResNet50 -- 2.3 InceptionV3 -- 3 Experimental Set-Up -- 3.1 Dataset -- 3.2 Free Parameters Specification -- 3.3 Performance of the Deep Learning Models -- 4 Results and Discussion -- 4.1 VGG16 Model. 4.2 ResNet50 Model -- 4.3 InceptionV3 Model -- 5 Conclusions and Future Research -- References -- Assessing GPT-4 Generated Abstracts: Text Relevance and Detectors Based on Faithfulness, Expressiveness, and Elegance Principle -- 1 Introduction -- 2 Literature Review -- 2.1 Comparison of GPT-4 to Other Large Language Models (LLMs) -- 2.2 The Evolution of GPT from GPT-1 to GPT-4 -- 2.3 Writing Capabilities in GPT-4 -- 3 Methods -- 3.1 Datasets -- 3.2 Prompting -- 3.3 Evaluation Metrics -- 4 Results -- 4.1 Text Relevance Scores -- 4.2 AI Detector Scores -- 4.3 Plagiarism Detector Scores -- 5 Discussion -- 6 Conclusion -- References -- Financial Text Sentiment Analysis Based on ChatGPT-Taking the Real Estate Industry as an Example -- 1 Introduction -- 2 Literature Review -- 2.1 Dictionary-Based Methods -- 2.2 Machine Learning-Based Methods -- 2.3 Emotion Lexicon - Machine Learning Combination Methods -- 2.4 Using ChatGPT Tool as an Analysis Method -- 3 Overview of ChatGPT Technology Process -- 4 Research Methodology -- 4.1 Workflow of the Experiments -- 4.2 Datasets -- 4.3 Preliminary Experiment: Industry Sentiment Training -- 4.4 Improvement Experiment: Enterprise Sentiment Training -- 5 Experiment and Results -- 5.1 Preliminary Experiment: Industry Sentiment Training -- 5.2 Improvement Experiment: Enterprise Sentiment Training -- 6 Conclusion and Future Work -- 6.1 Conclusion -- 6.2 Future Work -- References -- Reinforcement Learning Approaches -- A New Deep Reinforcement Learning Algorithm for UAV Swarm Confrontation Game -- 1 Introduction -- 2 Related Work -- 2.1 UAV Swarm Confrontation Strategy -- 2.2 Multi-agent Deep Reinforcement Learning -- 3 The UAV Swarm Confrontation Problem and the Algorithms -- 3.1 Task Description -- 3.2 Reward Function -- 3.3 The Implementation of MADDPG on the Task -- 3.4 The Description of MB-MADDPG. 4 Experiments and Analysis -- 4.1 Verification of Strategy Optimization Effect -- 4.2 Comparison and Analysis of Algorithms -- 5 Conclusion -- References -- Proximal Policy Optimization for Same-Day Delivery with Drones and Vehicles -- 1 Introduction -- 2 Problem Definition: SDDPHF -- 2.1 Problem Statement -- 2.2 Route-Based MDP -- 3 Solution Approach -- 4 Computational Results -- 4.1 Parameter Setting -- 4.2 Evaluation -- 5 Conclusion -- References -- A Double-Layer Reinforcement Learning Feature Optimization Framework for Evolutionary Computation Based Feature Selection Algorithms -- 1 Introduction -- 2 Background and Related Work -- 2.1 A Subsection Sample -- 2.2 Q-Learning -- 2.3 EC-Based Feature Selection Algorithms -- 2.4 Reinforcement Learning Based Feature Selection Algorithm -- 3 The Propose Method -- 3.1 Double-Layer Reinforcement Learning Based Feature Optimization Framework -- 3.2 Lower-Layer Feature Optimization Algorithm Based on Q-Learning --

3.3 Upper-Layer Length Control Strategic Based on Q-Learning -- 4  
Experiment and Result -- 4.1 Dataset Description and Experiment  
Settings -- 4.2 Parameter Settings -- 4.3 Experiment Result -- 5  
Conclusion -- References -- Combinatorial Optimization Approaches  
-- Ant-Antlion Optimizer with Similarity Information  
for Multidimensional Knapsack Problem -- 1 Introduction -- 2 Related  
Works -- 3 Ant-Antlion Optimizer with Similarity Information -- 3.1  
Novel Fitness with Similarity Information -- 3.2 Self-adaption Mutation  
-- 3.3 Procedure of Proposed Method -- 4 Experiments and Results --  
4.1 Testing Problems and Comparison Methods -- 4.2 Simulation  
and Evaluation -- 5 Conclusions -- References -- Efficient Graph  
Sequence Reinforcement Learning for Traveling Salesman Problem -- 1  
Introduction -- 2 Related Work -- 2.1 Autoregressive Methods -- 2.2  
Non-autoregressive Methods.  
3 Graph Sequence Reinforcement Learning Model.

---

Sommario/riassunto

This two-volume set, CCIS 2017 and 2018 constitutes the 8th International Conference, on Data Mining and Big Data, DMBD 2023, held in Sanya, China, in December 2023. The 38 full papers presented in this two-volume set included in this book were carefully reviewed and selected from 79 submissions. The papers present the latest research on advantages in theories, technologies, and applications in data mining and big data. The volume covers many aspects of data mining and big data as well as intelligent computing methods applied to all fields of computer science, machine learning, data mining and knowledge discovery, data science, etc.

---