

1. Record Nr.	UNINA9910851994303321
Autore	Yang De-Nian
Titolo	Advances in Knowledge Discovery and Data Mining : 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2024, Taipei, Taiwan, May 7–10, 2024, Proceedings, Part III // edited by De-Nian Yang, Xing Xie, Vincent S. Tseng, Jian Pei, Jen-Wei Huang, Jerry Chun-Wei Lin
Pubbl/distr/stampa	Singapore : , : Springer Nature Singapore : , : Imprint : Springer, , 2024
ISBN	9789819722594 9819722594
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (448 pages)
Collana	Lecture Notes in Artificial Intelligence, , 2945-9141 ; ; 14647
Altri autori (Persone)	XieXing TsengVincent S PeiJian HuangRenwei LinJerry Chun-Wei
Disciplina	006.3
Soggetti	Artificial intelligence Algorithms Education - Data processing Computer science - Mathematics Signal processing Computer networks Artificial Intelligence Design and Analysis of Algorithms Computers and Education Mathematics of Computing Signal, Speech and Image Processing Computer Communication Networks
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- General Chairs' Preface -- PC Chairs' Preface -- Organization -- Contents - Part III -- Interpretability and Explainability -- Neural

Additive and Basis Models with Feature Selection and Interactions -- 1 Introduction -- 2 Generalized Additive Models (GAMs) -- 2.1 Neural Additive Model (NAM) -- 2.2 Neural Basis Model (NBM) -- 3 NAM and NBM with Feature Selection -- 3.1 Motivation -- 3.2 Model Architecture -- 3.3 Implementation Remark -- 4 Discussion of Model Complexities -- 5 Experiments -- 5.1 Experimental Settings -- 5.2 Baselines -- 5.3 Results -- 6 Conclusion -- References -- Random Mask Perturbation Based Explainable Method of Graph Neural Networks -- 1 Introduction -- 2 Related Work -- 3 Problem Statement -- 4 Explainable Method -- 4.1 Node Importance Based on Fidelity -- 4.2 Explanation Sparsity -- 5 Experiments -- 5.1 Experimental Setup -- 5.2 Quantitative Experiments -- 5.3 Ablation Study -- 5.4 Use Case -- 6 Conclusion -- References -- RouteExplainer: An Explanation Framework for Vehicle Routing Problem -- 1 Introduction -- 2 Related Work -- 3 Proposed Framework: RouteExplainer -- 3.1 Many-to-Many Edge Classifier -- 3.2 Counterfactual Explanation for VRP -- 4 Experiments -- 4.1 Quantitative Evaluation of the Edge Classifier -- 4.2 Qualitative Evaluation of Generated Explanations -- 5 Conclusion and Future Work -- References -- On the Efficient Explanation of Outlier Detection Ensembles Through Shapley Values -- 1 Introduction -- 2 Related Work -- 3 Outlier Detection Ensembles -- 4 The bagged Shapley Values -- 5 Theoretical Guarantees for the Approximation -- 6 Experiments -- 6.1 Quality of the Approximation -- 6.2 Effectiveness -- 6.3 Scalability -- 7 Conclusions -- References -- Interpreting Pretrained Language Models via Concept Bottlenecks -- 1 Introduction -- 2 Related Work -- 2.1 Interpreting Pretrained Language Models. -- 2.2 Learning from Noisy Labels -- 3 Enable Concept Bottlenecks for PLMs -- 3.1 Problem Setup -- 4 C3M: A General Framework for Learning CBE-PLMs -- 4.1 ChatGPT-Guided Concept Augmentation -- 4.2 Learning from Noisy Concept Labels -- 5 Experiments -- 6 Conclusion -- A Definitions of Training Strategies -- B Details of the Manual Concept Annotation for the IMDB Dataset -- C Implementation Detail -- D Parameters and Notations -- E Statistics of Data Splits -- F Statistics of Concepts in Transformed Datasets -- G More Results on Explainable Predictions -- H A Case Study on Test-Time Intervention -- I Examples of Querying ChatGPT -- References -- Unmasking Dementia Detection by Masking Input Gradients: A JSM Approach to Model Interpretability and Precision -- 1 Introduction -- 2 Related Work -- 3 Methods -- 3.1 Jacobian Saliency Map (JSM) -- 3.2 Jacobian-Augmented Loss Function (JAL) -- 4 Experiments -- 4.1 Dataset -- 4.2 Preprocessing -- 4.3 Multimodal Classification -- 4.4 Performance Evaluation -- 5 Conclusion -- References -- Towards Nonparametric Topological Layers in Neural Networks -- 1 Introduction -- 1.1 Background -- 1.2 Motivation and Challenges -- 1.3 Contributions -- 2 Preliminaries and Related Work -- 2.1 Basics of Topology -- 2.2 Topological Neural Network -- 2.3 Functional Spaces for Machine Learning -- 3 Methodology -- 4 Evaluation -- 4.1 Experimental Setup -- 4.2 Implementation -- 4.3 Overall Performance -- 4.4 Learning Rate -- 4.5 Temporal-Spatial Correlation -- 5 Conclusion -- References -- Online, Streaming, Distributed Algorithms -- Streaming Fair k-Center Clustering over Massive Dataset with Performance Guarantee -- 1 Introduction -- 1.1 Problem Statement -- 1.2 Related Work -- 1.3 Our Contribution -- 2 A Two-Pass Algorithm with Approximation Ratio 3 -- 2.1 The ϵ -Independent Center Set -- 2.2 The Two-Pass Streaming Algorithm. -- 3 The Streaming Algorithm with an Approximation Ratio 7 -- 3.1 The Streaming Algorithm for Constructing 1 and 2 -- 3.2 Post-streaming Construction of Center Set C from 12 -- 4 Experimental Results -- 4.1

Experimental Setting -- 4.2 Experimental Analysis -- 5 Conclusion --
References -- Projection-Free Bandit Convex Optimization over
Strongly Convex Sets -- 1 Introduction -- 2 Related Work -- 2.1
Projection-Free OCO Algorithms -- 2.2 Bandit Convex Optimization --
3 Main Results -- 3.1 Preliminaries -- 3.2 Our Proposed Algorithm --
3.3 Theoretical Guarantees -- 4 Experiments -- 4.1 Problem Settings
-- 4.2 Experimental Results -- 5 Conclusion -- References -- Adaptive
Prediction Interval for Data Stream Regression -- 1 Introduction -- 2
Related Work -- 3 Background -- 4 Adaptive Prediction Interval(AdaPI)
-- 5 Experiments and Results -- 5.1 Comparison to Interval Forecast
-- 5.2 Comparison Between MVE and AdaPI -- 6 Conclusions --
References -- Probabilistic Guarantees of Stochastic Recursive Gradient
in Non-convex Finite Sum Problems -- 1 Introduction -- 1.1 Related
Works -- 1.2 Our Contributions -- 1.3 Notation -- 2 Prob-SARAH
Algorithm -- 3 Theoretical Results -- 3.1 Technical Assumptions --
3.2 Main Results on Complexity -- 3.3 Proof Sketch -- 4 Numerical
Experiments -- 4.1 Logistic Regression with Non-convex
Regularization -- 4.2 Two-Layer Neural Network -- 5 Conclusion --
References -- Rethinking Personalized Federated Learning with
Clustering-Based Dynamic Graph Propagation -- 1 Introduction -- 2
Related Work -- 3 Methodology -- 3.1 Model Overview -- 3.2 Client
Model Clustering -- 3.3 Dynamic Weighted Graph Construction -- 3.4
Knowledge Propagation and Aggregation -- 3.5 Precise Personalized
Model Distribution -- 4 Experiment -- 4.1 Experiment Setup -- 4.2
Performance Evaluation -- 4.3 Ablation Study -- 4.4 Case Study -- 4.5
Hyperparameter Study.
5 Conclusion -- References -- Unveiling Backdoor Risks Brought by
Foundation Models in Heterogeneous Federated Learning -- 1
Introduction -- 2 Related Work -- 3 Methodology -- 3.1 Threat Model
-- 3.2 FMs Empowered Backdoor Attacks to HFL -- 4 Experiment --
4.1 Experiment Setup -- 4.2 Experimental Results -- 4.3
Homogeneous Setting Evaluation -- 4.4 Case Study: Attack
Effectiveness v.s. Public Data Utilization Ratio -- 4.5 Hyper-Parameter
Study: ASR v.s. Poisoning Ratio -- 5 Conclusion -- References --
Combating Quality Distortion in Federated Learning with Collaborative
Data Selection -- 1 Introduction -- 2 Related Works -- 3 Proposal --
3.1 Preliminaries -- 3.2 Design Principle -- 3.3 Collaborative Sample
Selection (CSS) -- 4 Evaluation -- 4.1 Datasets and Experimental
Settings -- 4.2 Experimental Results -- 5 Conclusion -- References --
Probabilistic Models and Statistical Inference -- Neural Marked Hawkes
Process for Limit Order Book Modeling -- 1 Introduction -- 2
Background -- 3 Neural Marked Hawkes Process -- 4 Related Work --
5 Experiments -- 6 Conclusion -- References -- How Large Corpora
Sizes Influence the Distribution of Low Frequency Text n-grams -- 1
Introduction -- 2 Background and Related Work -- 3 The Model -- 4
Results -- 4.1 The Corpora Collection -- 4.2 The Range of k Values for
 $W(k,C -- L,n)$ Prediction -- 4.3 The Assessment Criteria and Parameter
Estimation -- 4.4 Comparison with Other Models -- 4.5 Obtained
Results -- 4.6 The Predictions with Growing Corpus Size -- 5
Conclusions -- References -- Meta-Reinforcement Learning Algorithm
Based on Reward and Dynamic Inference -- 1 Introduction -- 2
Background -- 2.1 Meta-Reinforcement Learning -- 2.2 Context-Based
Meta-Reinforcement Learning -- 2.3 Parametric Task Distributions -- 3
Problem Statement -- 4 Method -- 4.1 Reward and Dynamics
Inference.
4.2 Meta-Reinforcement Learning Algorithm Based on Reward and
Dynamics Inference Encoders -- 5 Experiment -- 5.1 Common MuJoCo
Environments -- 5.2 Cartesian Product Combinations of Tasks with

Different Goals and Dynamics -- 6 Discussion -- References -- Security and Privacy -- SecureBoost+: Large Scale and High-Performance Vertical Federated Gradient Boosting Decision Tree -- 1 Introduction -- 2 Preliminaries -- 2.1 Gradient Boosting Decision Tree -- 2.2 Paillier Homomorphic Encryption -- 2.3 SecureBoost -- 2.4 Performance Bottlenecks Analysis for SecureBoost -- 3 Proposed SecureBoost+ Framework -- 3.1 Ciphertext Operation Optimization -- 3.2 Training Mechanism Optimization -- 4 Experiments -- 4.1 Setup -- 4.2 Ciphertext Operation Optimization Evaluation -- 4.3 Training Mechanism Optimization Evaluation -- 5 Conclusion -- References -- Construct a Secure CNN Against Gradient Inversion Attack -- 1 Introduction -- 2 Preliminary -- 2.1 Federated Learning -- 2.2 Gradient Inversion Attack -- 2.3 Recursive Gradient Attack on Privacy (R-GAP) -- 3 Secure Convolutional Neural Networks -- 4 Experiment -- 4.1 Quantitative Results -- 4.2 Quantitative Results -- 5 Related Work -- 6 Limitation and Conclusion -- References -- Backdoor Attack Against One-Class Sequential Anomaly Detection Models -- 1 Introduction -- 2 Related Work -- 3 Preliminaries -- 3.1 Deep One-Class Sequential Anomaly Detection -- 3.2 Mutual Information Maximization -- 4 Methodology -- 4.1 Threat Model -- 4.2 The Proposed Attack -- 4.3 Post-deployment Attack -- 5 Experiments -- 5.1 Experimental Setup -- 5.2 Experimental Results -- 6 Conclusions -- References -- Semi-supervised and Unsupervised Learning -- DALLMi: Domain Adaption for LLM-Based Multi-label Classifier -- 1 Introduction -- 2 Language Model and Domain Adaptation -- 3 DALLMi -- 4 Experiments -- 5 Conclusion -- References. Contrastive Learning for Unsupervised Sentence Embedding with False Negative Calibration.

Sommario/riassunto

The 6-volume set LNAI 14645-14650 constitutes the proceedings of the 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2024, which took place in Taipei, Taiwan, during May 7–10, 2024. The 177 papers presented in these proceedings were carefully reviewed and selected from 720 submissions. They deal with new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, big data technologies, and foundations.
