Neurocomputing 470 (2022) 300–303

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Advances in artificial neural networks, machine learning and computational intelligence

Selected papers from the 28th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2020)

This special issue of Neurocomputing presents 13 original articles (2 of them are tutorials selected from the special sessions) that are extended versions of selected papers from the 29th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2020), a major event for researchers in the field of artificial neural networks and related topics. This single track conference is held annually in Bruges, Belgium, a UNESCO World Heritage Site and one of the most beautiful medieval city centers in Europe. It is jointly organized by UCL (Université Catholique de Louvain – Louvain-la-Neuve) and KU Leuven (Katholiek Universiteit – Leuven), and it is steered by Prof. Michel Verleysen from UCL. In addition to regular sessions, the conference regularly welcomes special sessions organized by renowned scientists in their respective fields. These sessions focus on specific topics, for instance deep learning, reinforcement learning, kernel methods, randomized learning approaches, biomedical applications, robotics, physics, environmental data analysis, clustering, data privacy, data visualization, and big data analytics.

The contributions in this special issue show that ESANN covers a broad range of topics in neural computing and neuroscience, from theoretical aspects to state-of-the-art applications. Around 200 researchers from 20 countries attended remotely (due to the COVID-19 pandemic) the 28th ESANN in October 2020. Around 100 oral and poster contributions have been presented. Based on the reviewers' and special session organisers' recommendations, as well as on the quality of the oral presentations at the conference, a number of authors were invited to submit an extended version of their conference paper for this special issue of Neurocomputing. All extended manuscripts went through an additional review process by at least two independent experts and the 13 articles presented in this volume were accepted for publication. They can be grouped as follows:

1. **Trustworthy Machine Learning**

   (a) André Artelt and Barbara Hammer*Efficient computation of counterfactual explanations and counterfactual metrics of prototype-based classifiers*The increasing use of machine learning in practice and legal regulations like EU's GDPR cause the necessity to be able to explain the prediction and behaviour of machine learning models. A prominent example of particularly intuitive explanations of AI models in the context of decision making are counterfactual explanations. Yet, it is still an open research problem how to efficiently compute counterfactual explanations for many models. In this contribution, authors investigate how to efficiently compute counterfactual explanations for an important class of models, prototype-based classifiers such as learning vector quantisation models.In particular, authors derive specific convex and non-convex programs depending on the used metric.Typically counterfactual explanations deliver a feedback in terms of changes of the input features which lead to a different output - one application scenario is the link of these required changes to actionable items to change the desired outcome. Yet, rather than minimum changes of the input, it is interesting to address minimum changes of the model itself, which are required to lead to a different result rather than a change of its inputs. Authors phrase this question as a counterfactual of the model prescription rather than the data points. Authors focus on distance-based classifiers (in particular learning vector quantisation models), where model changes correspond to changes of metric parameters, and they develop efficient optimisation techniques to generate such counterfactual metric changes depending on the chosen model.

   (b) Danilo Franco, Nicolò Navarin, Michele Donini, Davide Anguita, and Luca Oneto*Deep Fair Models for Complex Data: Graphs Labelling and Explainable Face Recognition*The central goal of algorithmic fairness is to develop AI-based systems which do not discriminate subgroups in the population with respect to one or multiple notions of inequity, knowing that data is often humanly biased. Researchers are racing to develop AI-based systems able to reach superior performance in terms of accuracy, increasing the risk of inheriting the human biases hidden in the data. An obvious tension exists between these two lines of research that are currently colliding due to increasing concerns regarding the widespread adoption of these systems and their ethical impact. The problem is even more challenging when the input data is complex (e.g., graphs, trees, or images) and deep uninterpretable models need to be employed to achieve satisfactory performance. In fact, it is required to develop a deep architecture to learn a data representation able, from one side, to be expressive enough to describe the data and lead to highly accurate models and, from the other side, to discard

https://doi.org/10.1016/j.neucom.2021.07.053

all the information which may lead to unfair behaviour. In this work authors measure fairness according to Demographic Parity, requiring the probability of the model decisions to be independent of the sensitive information. Authors investigate how to impose this constraint in the different layers of deep neural networks for complex data, with particular reference to deep networks for graph and face recognition. Authors present experiments on different real-world datasets, showing the effectiveness of their proposal both quantitatively by means of accuracy and fairness metrics and qualitatively by means of visual explanation.

(c) Julia Lust and Alexandru P. Condurache*Efficient Detection of Adversarial, Out-of-distribution and Other Misclassified Samples*Deep Neural Networks (DNNs) are increasingly being considered for safety–critical approaches in which it is crucial to detect misclassified samples. Typically, detection methods are geared towards either the detection of out-of-distribution or adversarial data. Additionally, most detection methods require a significant amount of parameters and runtime. In this contribution authors discuss a novel approach for detecting misclassified samples suitable for out-of-distribution, adversarial and additionally real world error-causing corruptions. It is based on the Gradient's Norm (GraN) of the DNN and is parameter and runtime efficient. Authors evaluate GraN on two different classification DNNs (DenseNet, ResNet) trained on different datasets (CIFAR-10, CIFAR-100, SVHN). In addition to the detection of different adversarial example types (FGSM, BIM, Deepfool, CWL2) and out-of-distribution data (TinyImageNet, LSUN, CIFAR-10, SVHN) authors evaluate GraN for novel corruption setups (Gaussian, Shot and Impulse noise). Authors experiments show that GraN performs comparable to state-of-the-art methods for adversarial and out-of-distribution detection and is superior for real world corruptions while being parameter and runtime efficient.

(d) Jan Philip Göpfert, Heiko Wersing, and Barbara Hammer*Interpretable Locally Adaptive Nearest Neighbors*When training automated systems, it has been shown to be beneficial to adapt the representation of data by learning a problem-specific metric. This metric is global. Authors extend this idea and, for the widely used family of k nearest neighbours algorithms, develop a method that allows learning locally adaptive metrics. These local metrics not only improve performance, but are naturally interpretable. To demonstrate important aspects of how authors approach works, they conduct a number of experiments on synthetic data sets, and they show its usefulness on real-world benchmark data sets.

2. **Efficient and Accurate Machine Learning**

(a) Benjamin Paaßen, Alexander Schulz, and Barbara Hammer*Reservoir Stack Machines*Memory-augmented neural networks equip a recurrent neural network with an explicit memory to support tasks that require information storage without interference over long times. A key motivation for such research is to perform classic computation tasks, such as parsing. However, memory-augmented neural networks are notoriously hard to train, requiring many backpropagation epochs and lots of data samples. In this paper, authors introduce the reservoir stack machine, a model which can provably recognise all deterministic context-free languages and circumvents the training problem by training only the output layer of a recurrent net and employing auxiliary information during training about the desired interaction with a stack. In authors experiments, they validate the reservoir stack machine against deep and shallow networks from the literature on three benchmark tasks for Neural Turing machines and six deterministic context-free languages. Authors results show that the reservoir stack machine achieves zero error, even on test sequences longer than the training data, requiring only a few seconds of training time and 100 training sequences.

(b) Laura Morán-Fernández, and Veronica Bolón-Canedo, Amparo Alonso-Betanzos*How important is data quality? Best classifiers vs best features*The task of choosing the appropriate classifier for a given scenario is not an easy-to-solve one. First, there is an increasingly high number of algorithms available belonging to different families. And also there is a lack of methodologies that can help on recommending in advance a given family of algorithms for a certain type of datasets. Besides, most of these classification algorithms exhibit a degradation in the performance when faced with datasets containing irrelevant and/or redundant features. In this work authors analyse the impact of feature selection in classification over several synthetic and real datasets. The experimental results obtained show that the significance of selecting a classifier decreases after applying an appropriate pre-processing step and, not only this alleviates the choice, but it also improves the results in almost all the datasets tested.

(c) Abolfazl Taghribi, Marco Canducci, Michele Mastropietro, Sven De Rijcke, Kerstin Bunte, and Peter Tiňo*ASAP - A Sub-sampling Approach for Preserving Topological Structures Modelled with Geodesic Topographic Mapping*Topological data analysis tools enjoy increasing popularity in a wide range of applications, such as Computer graphics, Image analysis, Machine learning, and Astronomy for extracting information. However, due to computational complexity, processing large numbers of samples of higher dimensionality quickly becomes infeasible. The contribution of this work is twofold: authors present an efficient novel sub-sampling strategy inspired by Coulomb's law to decrease the number of data points in d-dimensional point clouds while preserving its homology. The method is not only capable of reducing the memory and computation time needed for the construction of different types of simplicial complexes but also preserves the size of the voids in d-dimensions, which is crucial for example for astronomical applications. Furthermore, authors propose a technique to construct a probabilistic description of the border of significant cycles and cavities inside the point cloud. Authors demonstrate and empirically compare the strategy in several synthetic scenarios and an astronomical particle simulation of a dwarf galaxy for the detection of superbubbles (supernova signatures).

3. **Machine Learning in Structured Domains**

(a) Filippo Maria Bianchi, Claudio Gallicchio, and Alessio Micheli*Pyramidal Reservoir Graph Neural Network*Authors propose a deep Graph Neural Network (GNN) model that alternates two types of layers. The first type is inspired by Reservoir Computing (RC) and generates new vertex features by iterating a non-linear map until it converges to a fixed point. The second type of layer implements graph pooling operations, that gradually reduce the support graph and the vertex features, and further improve the computational efficiency of the RC-based GNN. The architecture is, therefore, pyramidal. In the last layer, the features of the remaining vertices are combined into a single vector, which represents the graph embedding. Through a mathematical derivation introduced in this paper, authors show formally how graph pooling can reduce the computational complexity of the model and speed-up the convergence of the dynamical updates of the vertex features. The proposed approach to the design of RC-based GNNs offers an advanta-

geous and principled trade-off between accuracy and complexity, which they extensively demonstrate in experiments on a large set of graph datasets.

(b) Daniele Castellana and Davide Bacciu*A Tensor Framework for Learning in Structured Domains*Learning machines for structured data, such as trees, are intrinsically based on their capacity to learn representations by aggregating information from the multi-way relationships emerging from the structure topology. Complex aggregation functions are desirable in this context to increase the expressiveness of the learned representations. Unfortunately, the modelling of higher-order interactions among structure constituents is practically unfeasible due to the exponential number of parameters required. Therefore, the common approach is to define models which rely only on first-order interactions among structure constituents. In this work, authors leverage tensors theory to define a framework for learning in structured domains. Such a framework is built on the observation that more expressive models require a tensor parameterisation. This observation is the stepping stone for the application of tensor decompositions in the context of recursive models. From this point of view, the advantage of using tensor decompositions is twofold since it allows limiting the number of model parameters while injecting inductive biases that do not ignore higher-order interactions. Authors apply the proposed framework on probabilistic and neural models for structured data, defining different models which leverage tensor decompositions. The experimental validation clearly shows the advantage of these models compared to first-order and full-tensorial models.

4. **Machine Learning in Bioinformatics**

(a) Thomas Beznik, Paul Smyth, Gael de Lannoy, and John A. Lee*Deep Learning to Detect Bacterial Colonies for the Production of Vaccines* During the development of vaccines, bacterial colony forming units (CFUs) are counted in order to quantify the yield in the fermentation process. This manual task is long, tedious, and subject to errors. In this work, multiple segmentation algorithms based on the U-Net CNN architecture are tested and proven to offer robust, automated CFU counting. It is also shown that the multiclass generalisation with a bespoke loss function allows virulent and avirulent colonies to be distinguished with acceptable accuracy. While many possibilities are left to explore, their results show the potential of deep learning for separating and classifying bacterial colonies.

(b) Kiki van der Heijden and Siamak Mehrkanoon*Goal-driven, neurobiological-inspired convolutional neural network models of human spatial hearing*The human brain effortlessly solves the complex computational task of sound localisation using a mixture of spatial cues. How the brain performs this task in naturalistic listening environments, for example with reverberation, is not well understood. In the paper the authors build on the success of deep neural networks at solving complex and high-dimensional problems to develop goal-driven, neurobiological-inspired convolutional neural network (CNN) models of human spatial hearing. After training, authors visualise and quantify feature representations in intermediate layers to gain insights into the representational mechanisms underlying sound location encoding in CNNs. Authors results show that neurobiological-inspired CNN models trained on real-life sounds specialised with human binaural hearing characteristics can accurately predict sound location in the horizontal plane. CNN localisation acuity across the azimuth resembles human sound localisation acuity, but CNN models outperform human sound localisation in the back. Training models with different objective functions - that is, minimising either Euclidean or angular distance - modulates localisation acuity in particular ways. Moreover, different implementations of binaural integration result in unique patterns of localisation errors that are in agreement with behavioural observations in humans. Finally, feature representations reveal a gradient of spatial selectivity across network layers, starting with broad spatial representations in early layers and progressing to sparse, highly selective spatial representations in deeper layers. In summary the authors results show that neurobiological-inspired CNNs are a valid approach to modelling human spatial hearing, Moreover, results pave the way for future studies combining neural network models with empirical measurements of neural activity to unravel the complex computational mechanisms underlying neural sound location encoding in the human auditory pathway.

5. **Tutorials**

(a) Ivano Lauriola, Alberto Lavelli, and Fabio Aiolli*An introduction to Deep Learning in Natural Language Processing: models, techniques, and tools*Natural Language Processing (NLP) is a branch of artificial intelligence that involves the design and implementation of systems and algorithms able to interact through human language. Thanks to the recent advances of deep learning, NLP applications have received an unprecedented boost in performance. In this paper, authors present a survey of the application of deep learning techniques in NLP, with a focus on the various tasks where deep learning is demonstrating stronger impact. Additionally, authors explore, describe, and revise the main resources in NLP research, including software, hardware, and popular corpora. Finally, authors emphasise the main limits of deep learning in NLP and current research directions.

(b) José D. Martín-Guerrero, and Lucas Lamata*Quantum Machine Learning: A tutorial*This tutorial provides an overview of Quantum Machine Learning (QML), a relatively novel discipline that brings together concepts from Machine Learning (ML), Quantum Computing (QC) and Quantum Information (QI). The great development experienced by QC, partly due to the involvement of giant technological companies as well as the popularity and success of ML have been responsible of making QML one of the main streams for researchers working on fuzzy borders between Physics, Mathematics and Computer Science. A possible, although arguably coarse, classification of QML methods may be based on those approaches that make use of ML in a quantum experimentation environment and those others that take advantage of QC and QI to find out alternative and enhanced solutions to problems driven by data, oftentimes offering a considerable speedup and improved performances as a result of tackling problems from a complete different standpoint. Several examples are provided to illustrate both classes of methods.

The guest editors would like to thank all authors for their interesting contributions and all reviewers for their excellent work. Authors and reviewers were asked to respect a very tight schedule, which allowed this issue to be published in less than a year after the conference, timely before the ESANN meeting of 2021. We would also like to thank the Neurocomputing editorial board for giving us the opportunity to publish this issue, as well as Elsevier's people for the very efficient and seamless management of the publication procedure. Finally, our most sincere gratitude goes to Prof. Michel Verleysen for his excellent conference organisation, and his strong support to the realisation of this special issue.

Luca Oneto was born in Rapallo, Italy in 1986. He received his BSc and MSc in Electronic Engineering at the University of Genoa, Italy respectively in 2008 and 2010. In 2014 he received his PhD

from the same university in the School of Sciences and Technologies for Knowledge and Information Retrieval with the thesis "Learning Based On Empirical Data". In 2017 he obtained the Italian National Scientific Qualification for the role of Associate Professor in Computer Engineering and in 2018 he obtained the one in Computer Science. He worked as Assistant Professor in Computer Engineering at University of Genoa from 2016 to 2019. In 2018 he was co-funder of the spin-off ZenaByte s.r.l. In 2019 he obtained the Italian National Scientific Qualification for the role of Full Professor in Computer Science and Computer Engineering. In 2019 he became Associate Professor in Computer Science at University of Pisa and currently is Associate Professor in Computer Engineering at University of Genoa. He has been involved in several H2020 projects (S2RJU, ICT, DS) and he has been awarded with the Amazon AWS Machine Learning and Somalvico (best Italian young AI researcher) Awards. His first main topic of research is the Statistical Learning Theory with particular focus on the theoretical aspects of the problems of (Semi) Supervised Model Selection and Error Estimation. His second main topic of research is Data Science with particular reference to the problem of Trustworthy AI and the solution of real world problems by exploiting and improving the most recent Learning Algorithms and Theoretical Results in the fields of Machine Learning and Data Mining. Further information can be found at www.lucaoneto.it

Kerstin Bunte graduated at the Faculty of Technology at the University of Bielefeld, Germany. Since July 2016 she is assistant professor and Rosalind Franklin Fellow at the University of Groningen. Dr. Bunte received her Ph.D. in Computer Science in 2011 from the University of Groningen. She had research visits at the University of Rochester (USA) and postdoctoral positions have taken her to Bielefeld University (Germany), Aalto University (Finland), and Université catholique de Louvain (Belgium). In 2015 she got a European Marie Skłodowska-Curie Fellowship (Project ID: 659104) at the University of Birmingham and she was partner in the European ITN: SUNDIAL, Project ID: 721463 and the Regions of Smart Factories project. Since 2020 she is project leader of the NWO Smart Industry project SMART-AGENTS and recieved an NWO VIDI grant for mechanistic machine learning. Her recent work has focused on the development of interpretable machine learning techniques (explainable AI) for interdisciplinary data analysis, dimensionality reduction and visualization. Further information can be found at www.cs.rug.nl/ kbunte

Nicoló **Navarin** is Assistant Professor in Computer Science at the Department of Mathematics "Tullio Levi–Civita", University of Padua, Italy. He received his B.Sc. (2008) and M.Sc. (2010) in Computer Science from the University of Padua. In 2014, he received the PhD in Computer Science from the University of Bologna. He has been a visiting researcher at the University of Freiburg, Germany, and a research fellow at the University of Nottingham, UK and at the University of Padua. His research interests are in the field of machine learning, including kernel methods and neural networks for structured data, and applications to bioinformatics, business process mining, computer vision and computational psychology. Dr. Navarin has been actively involved in the organisation of several machine learning conferences. He is a member of IEEE Computational Intelligence Society, IEEE Task Force on Deep learning, IEEE Task Force on Learning from Structured Data, and of the Italian Association for Artificial Intelligence. Further information can be found at www.math.unipd.it/nnavarin

Luca Oneto *
DIBRIS, Università di Genova, Italy
* Corresponding author.

Kerstin Bunte
University of Groningen, Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, The Netherlands

Nicoló Navarin
Dipartimento di Matematica "Tullio Levi-Civita", Università di Padova, Italy

Received 1 July 2021
Accepted 17 July 2021

Available online 24 July 2021