



Engineering Studies

Issue 3 (2), 2018

VOLUME 10

Taylor & Francis

2018

Engineering Studies, Issue 3 (2), Volume 10. Taylor & Francis, 2018. - Pages 300-832.

The edition materials are posted in Scopus and Web of Science.

Source Normalized Impact per Paper (SNIP): 0.514

SCImago Journal Rank (SJR): 0.217

Impact Factor: 0.234

*2016 Journal Citation Reports®, Thomson Reuters

Editor-in-Chief

Cyrus Mody - *Chair in the History of Science, Technology and Innovation, University of Maastrich, The Netherlands*

Deputy Editor:

Kacey Beddoes - *Department of Sociology, University of Massachusetts Lowell, USA*

Founding Editor:

Gary Downey - *Department of Science and Technology in Society, Virginia Tech, Blacksburg, Virginia, USA*

Associate Editors:

Atsushi Akeru - *Department of Science & Technology Studies, Rensselaer Polytechnic Institute, USA*

Konstantinos Chatzis - *University Paris-Est - LATTIS (UMR CNRS), France*

Maria Paula Diogo - *Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal*

Aditya Johri - *Department of Applied Information Technology, George Mason University, USA*

Ulrik Jørgensen - *Department of Development and Planning, Aalborg University, Denmark*

Vivian Anette Lagesen - *Department of Interdisciplinary Studies of Culture, Norwegian University of Science and Technology, Norway*

Donna Riley - *School of Engineering Education, Purdue University, USA*

Matthew Wisnioski - *Department of Science and Technology in Society, Virginia Tech, USA*

Chyuan-Yuan Wu - *Institute of Sociology and STS Program, National Tsing Hua University, Taiwan*

Web Editor

Brent Jesiek - *School of Engineering Education and School of Electrical and Computer Engineering, Purdue University, USA*

Assistant Editors

Andrew Chilvers - *Department of Civil, Environmental and Geomatic Engineering, University College London, UK*

Dana Denick - *National Science Foundation, Division of Civil, Mechanical, and Manufacturing Innovation, USA*

Ida Ngambeki - *Department of Computer and Information Technology, Purdue Polytechnic Institute, USA*

Justin Hess - *STEM Education Research Institute (SERI), Indiana University-Purdue University Indianapolis, USA*

Editorial Advisory Board

Stephen R. Barley - *Department of Management Science and Engineering, Stanford University, USA*

Sharon Beder - *Science, Technology & Society Program, University of Wollongong, Australia*

Bruno Belhoste - *Centre d'histoire des sciences et d'histoire des techniques, Université Paris 1 Panthéon-Sorbonne, France*

Li Bocong - *Department of Social Sciences, Chinese Academy of Sciences, China*

Taft Broome - *Department of Civil Engineering, Howard University, USA*

Louis Bucciarelli - *Program in Science, Technology and Society, Massachusetts Institute of Technology, USA*
Ivan da Costa Marques - *Graduate Program in Informatics, Universidade Federal do Rio de Janeiro, Brazil*
Michael Davis - *Department of Humanities, Illinois Institute of Technology, USA*
Christelle Didier - *Département d'éthique, Université Catholique de Lille, France*
Wendy Faulkner - *Science Studies Unit, University of Edinburgh, UK*
David E. Goldberg - *Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, USA*
Irina Gouzevitch - *Centre Alexandre Koyé, Museum National d'Histoire Naturelle, France*
André Grelon - *L'Ecole des Hautes Etudes en Sciences Sociales, France*
Deborah Johnson - *Department of Science, Technology, and Society, University of Virginia, USA*
Ronald Kline - *Departments of Electrical Engineering and Science & Technology Studies, Cornell University, USA*
Eda Kranakis - *Department of History, University of Ottawa, Canada*
Gideon Kunda - *Department of Labor Studies, Tel Aviv University, Israel*
Jang Gyu Lee - *School of Electrical Engineering and Computer Science, Seoul National University, Korea*
Juan Lucena - *Liberal Arts and International Studies, Colorado School of Mines, USA*
Peter Meiksins - *Department of Sociology, Cleveland State University, USA*
Carl Mitcham - *Liberal Arts and International Studies, Colorado School of Mines, USA*
Antoine Picon - *Department of Architecture, Harvard University, USA*
Bruce Seely - *Department of Social Sciences, Michigan Technological University, USA*
Sheri Sheppard - *Department of Mechanical Engineering, Stanford University, USA*
Amy Slaton - *Department of History and Politics, Drexel University, USA*
Knut H. Sørensen - *Department of Interdisciplinary Studies of Culture, Norwegian University of Science and Technology, Norway*
Dominique Vinck - *LADHUL, Institute of Social Sciences, University of Lausanne, Switzerland*
Rosalind Williams - *Program in Science, Technology and Society, Massachusetts Institute of Technology, USA*

Print ISSN: 1937-8629
Online ISSN: 1940-8374

© Taylor & Francis, 2018

CONTENTS

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Influence of NiCl₂/CdCl₂ as Mixed Filler on Structural, Thermal and Electrical Properties of PVA/PVP Blend <i>I.S. Elashmawi, E.M. Abdelrazek, A.Y. Yassin</i> | 307 |
| A Genetic Algorithm for Optimizing Background Subtraction Parameters in Computer Vision <i>Ramesh Rajagopalan</i> | 327 |
| Novel Data Mining Techniques for Incomplete Clinical Data in Diabetes Management <i>Herbert F. Jelinek, Andrew Yatsko, Andrew Stranieri, Sitalakshmi Venkatraman</i> | 337 |
| Anti-Inflammatory Effects of New Flavonoids from <i>Streptomyces</i> sp. BT01 in Lipopolysaccharide-Stimulated RAW 264.7 Murine Macrophages via Inhibition of NF-KappaB Activation <i>Thongchai Taechowisan, Srisakul Chanaphat, Wanwika Ruensamran, Waya S. Phutdhawong</i> .. | 359 |
| Nonlinear Programming in Precast Production and Transportation Planning <i>Aleksander Nicał</i> | 385 |
| Description of Vibration and Testing of the Water Turbine <i>Bogdan Żółtowski, Mariusz Żółtowski</i> | 395 |
| Real-time Power Calculation Using Lissajous Curve Method for DBD Reactors <i>Jacek Kołek, Tomasz Jakubowski</i> | 406 |
| Farm Tractor Reliability Quantification with the DPO Indicator <i>Karol Durczak, Mirosław Czechłowski, Adam Ekielski, Piotr Jurek</i> | 417 |
| Dual-diode Model of Functioning of Photovoltaic Modules in MATLAB Package <i>Mariusz Sarniak</i> | 437 |
| The Problem of Considering of the Torsional Stiffness of Reinforced Concrete Elements While Designing Statically Indeterminate Systems <i>Taliat Azizov, Nadzieja Jurkowska</i> | 453 |
| Optimization and Computational Intelligence Methods in Control Engineering Education of Marine Officers <i>Józef Lisowski</i> | 467 |
| About Systematic Errors in Measurements Microprofile Automobile Roads <i>A.M. Menshikov</i> | 481 |
| Use of Elderberry Juice for the Preparation of Soft Drinks <i>V. Koshova, A. Kobernitska</i> | 499 |
| The Analysis of the Accidents Causes at the Different Stages of the Construction Object's Life Cycle <i>Anastaia Lapina, Kseniya Shentsova, Andrey Ponomarenko</i> | 508 |

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Hybrid Method for the Big Data Analysis Using Neural Networks <i>Siranush Sargsyan, Anna Hovakimyan, Manya Ziroyan</i> | 519 |
| The Environment Work as a Basis for Sustainable Development of the Region <i>E.G. Meshchaninova, O.A. Tkacheva</i> | 527 |
| Method of Measurement of Interference in the Chains of the On-board Equipment of Spacecrafts and Processing Their Results <i>A.V. Kostin, M.N. Piganov, V.S. Bozrikov</i> | 536 |
| The Use of Aromaticroot Vegetables in Technology of Freshwater Fish Preserves Production <i>Natalia Golembovskaya, Natalia Slobodyanuk, Tetyana Lebskaya</i> | 544 |
| Technique of Determining the Set of Informative Parameters for the Personal Prediction of the Quality and Reliability of Radio-Electronic Means <i>Roman Mishanov, Mikhail Piganov</i> | 565 |
| The Creation of Three-dimensional Model of a Relief with Application of Programs Digitals and Robur <i>O.V. Germak, O.A. Gugueva, N.A. Kalacheva</i> | 576 |
| Distribution of Information System Resources in Monitoring the Flows of Random Processes <i>Nikolay Boldyrikin, Viktoria Boldyrihina, Anastasia Manakova, Olga Safaryan</i> | 586 |
| Anti-Landslide Engineering Measures and Their Role in Reducing the Probability of Occurrence of Exogenous Geological Processes in the Chechen Republic <i>R.A. Gakaev</i> | 596 |
| Determination of Spatial Objects Position via Industrial Robot Systems <i>Rahim Gurban Mammadov, Timur Chingiz Aliyev</i> | 604 |
| The Stress-Strain State of the Two Shtreks in a Weighty Anisotropic Solid <i>Nurlan Azhikhanov, Tursinbay Turymbetov, Madenov Almat, Zhenis Aimeshov</i> | 613 |
| Engineering Geological Conditions of Karst Formation in the Chechen Republic <i>R.A. Gakaev</i> | 619 |
| Comparison of Two Criteria for Evaluating the Efficiency of Short-term Runoff Forecasting Methods <i>Edgar Fabian Espitia Sarmiento, Natalia V. Victorova, Vitaly A. Khaustov, Ekaterina V. Gaidukova</i> | 626 |
| Intelligent Control System of Indoor Microclimate Formation Process <i>B.B. Orazbayev, L.Zh. Sansyzbai, B.E. Utenova, L.T. Kurmangazieva</i> | 634 |
| Refinement Method Needs Spare Parts for Technical Service of Machine <i>Anna Anosova, Mikhail Buraev, Stepan Lugovnin, Alexey Shisteev</i> | 646 |
| Multidetails Clothes Design with Rigid Textile and Nontextile Details <i>Mariya Makarevich, Ekaterina Lunina</i> | 653 |

Hybrid Method for the Big Data Analysis Using Neural Networks

Siranush Sargsyan,

Yerevan State University, Armenia, Associate Professor, Ph.D.,

Chair of Programming and Information Technologies,

Anna Hovakimyan,

Yerevan State University, Armenia, Associate Professor, Ph.D.,

Chair of System Programming,

Manya Ziroyan,

Russian State Social University, Russia, Professor, Ph.D.,

Chair of Applied Mathematics and Informatics

Abstract: The large accumulation of data leads to a continuous increase in their volume, and the tasks for analyzing such data, solved by previous means, differ markedly and become more complicated. Appropriate information technologies, which give analysts the opportunity to study Big Data, should help in converting the information embedded in the data into new knowledge. Practical methods and means of supporting the high speed of data acquisition and analysis are of practical interest.

In the paper a hybrid method is proposed for analyzing Big Data in real time using the adaptive clustering method and the neural network model developed by the authors is proposed.

A classification model was developed and implemented using statistical methods and neural networks. This parametric system, analyzing the input data, establishes the appropriate parameters and, based on the task at hand, selects the solution method and implementation language. For the selection of system parameters, a large number of typical examples were constructed on which the system was tested.

Keywords: clusterization, classification, Big Data, neural network, real-time data analysis.

***Гибридный метод анализа больших данных
с помощью нейронных сетей***

Саркисян Сирануш,

*Ереванский государственный университет, Армения
доцент, кандидат физ.-мат. наук, кафедра программирования
и информационных технологий,*

Овакимян Анна,

*Ереванский государственный университет, Армения,
доцент, кандидат техн. наук, кафедра системного программирования,*

Зироян Маня,

*Российский государственный социальный университет, Россия, профессор,
доктор эк. наук, кафедра прикладной математики и информатики*

Аннотация: Постоянное накопление данных приводит к непрерывному росту их объема, и задачи для анализа таких данных, решаемые прежними средствами, заметно отличаются и усложняются. Соответствующие информационные технологии, дающие аналитикам возможность изучать большие данные, должны помогать в преобразовании заложенной в данных информации в новое знание. При этом практический интерес представляют методы и средства поддержки высокой скорости получения и анализа данных.

В работе предложен гибридный метод анализа больших данных в режиме реального времени с использованием метода адаптивной кластеризации и разработанной авторами нейросетевой модели.

Разработана и реализована модель классификации с использованием статистических методов и нейронных сетей. Эта параметрическая система, анализируя входные данные, устанавливает соответствующие параметры и, исходя из поставленной задачи, выбирает метод решения и язык реализации. Для подбора параметров системы было построено большое количество типичных примеров, на которых система была протестирована.

Ключевые слова: кластеризация, классификация, большие данные, нейронная сеть, анализ данных в реальном времени.

Введение

Накопление больших объемов данных в последнее время сделали актуальными задачи, предназначенные для извлечения, сбора и представления ко-

нечному пользователю информации, необходимой для их дальнейшего анализа и прогнозирования в будущем.

При наличии разных технологий, методов, средств и программных продуктов актуальной остается разработка и исследование новых методов, используемых в системах интеллектуального анализа данных (ИАД). Анализ большого объема данных требует особого подхода, подготовки исходных данных, специализированных алгоритмов и построения разных сценариев обработки данных.

Одной из эффективных стратегией обработки больших данных является разбиение данных на сегменты и построение моделей для каждого сегмента по отдельности. Подобный подход помогает повышению скорости анализа благодаря обработке меньших объемов данных за один проход. В этом случае отдельные обработки можно распараллелить, экономя затраченное время. К тому же модели для каждого сегмента могут строиться разными способами [1].

Настоящая работа посвящена разработке и исследованию метода создания модели для анализа данных, которая использует алгоритмы кластеризации и нейронные сети [2, 4, 5]. При анализе больших объемов данных на первом этапе решается задача кластеризации, то есть группировка данных по классам (кластерам). Каждый класс снабжается специфичной моделью анализа данных. На втором этапе создается модель нейронной сети, для обучения которой определяется репрезентативная выборка объектов из разных кластеров, и организуется ее обучение. После инициализации модели нейронной сети решается задача классификации объектов в реальном времени (Real-Time Analytics).

1. Интеллектуальный анализ данных

Интеллектуальный анализ данных связан с поиском в данных скрытых нетривиальных и полезных закономерностей, позволяющих получить новые знания об исследуемых данных. Так как постоянное накопление данных приводит к непрерывному росту их объема, известные статистические методы покрывают лишь часть нужд по обработке этих данных. В такой ситуации методы ИАД приобретают особую актуальность. Их основная особенность заключается в установлении наличия и характера скрытых закономерностей в данных, тогда как традиционные методы занимаются главным образом параметрической оценкой уже установленных закономерностей. Интеллектуальная система осу-

ществляет поиск функциональных и логических закономерностей в накопленных данных, построение моделей и правил, которые объясняют найденные закономерности и прогнозируют развитие некоторых процессов. Методы Data Mining помогают решить многие задачи, с которыми сталкивается аналитик. Среди методов ИАД особое место занимают классификация (Classification) и кластеризация (Clustering) [2-6].

В результате решения задачи классификации обнаруживаются признаки, по которым новый объект можно отнести к тому или иному классу. Основные проблемы, с которыми сталкиваются при решении задач классификации - это неудовлетворительное качество исходных данных, в которых встречаются как ошибочные данные, так и пропущенные значения.

Кластеризация, основываясь на отношении схожести элементов, формирует подмножества (кластеры), в которые группируются входные данные. Существует много методов и алгоритмов кластеризации, широко используемых в системах ИАД. Особый интерес к методам анализа данных возник в связи с развитием средств сбора и хранения данных, позволивших накапливать большие объемы информации. Нужно было от методов обработки собранных данных переходить к методам превращения их в знания.

Особую актуальность имеет разработка методики адаптивной кластеризации, отличительной особенностью которой является оценка качества каждого разбиения и выбор наилучшего из них [2]. Практическая значимость этой методики состоит в том, что метод позволяет на базе нечеткого отношения эквивалентности эффективно выявлять в данных кластеры произвольной формы. Методика адаптивной кластеризации включает предварительную подготовку и нормирование данных, определение целей анализа, определение средств анализа и способа представления результатов, выбор критериев оценки качества решения, анализ результатов. Методика адаптивной кластеризации и критерии оценки качества решения задачи кластеризации нашли свое применение в рамках реализации многоагентной системы интеллектуального анализа [2, 3].

2. Применение нейронных сетей в задачах ИАД

Искусственные нейронные сети прочно вошли в нашу жизнь и в настоящее время широко используются при решении самых разных задач и активно применяются там, где обычные алгоритмические решения оказываются неэф-

эффективными или вовсе невозможными. Нейросетевой подход используется для выявления линейных и нелинейных зависимостей между переменными и весьма эффективен. Нейронные сети, в отличие от статистических методов параметрического анализа, базируются на параллельной обработке информации и могут работать с неполными данными [5, 6,8].

Сегодня существует большое число различных конфигураций нейронных сетей с различными принципами функционирования. Способность к моделированию нелинейных процессов, работа с зашумленными данными и адаптивность дают возможность применять нейронные сети для решения широкого класса задач [3, 5-8].

Нейронная сеть – параллельно распределенная структура обработки информации, состоящая из нейронов, которые соединены между собой. Организация нейронов и их связей в определенную структуру (архитектуру) оказывает значительное влияние на вычислительные возможности нейронной сети [6, 8]. Для решения задачи нейросетевую модель следует обучить. Формирование обучающей выборки, которая должна удовлетворять требованиям полноты и непротиворечивости, является одной из основных задач при построении нейросетевой модели. Для проверки обучающей выборки на соответствие этим требованиям необходима оценка ее качества, разработка качественных критериев и математического аппарата для их однозначной оценки [5, 6]. Обученная модель проверяется на тестовых данных, то есть проверяется качество работы модели, и, если оно неудовлетворительно, происходит дополнительное ее обучение. Так продолжается до тех пор, пока не будет достигнут требуемый уровень качества. Затем модель применяют к анализируемым объектам и определяется, к каким из известных классов относятся исследуемые объекты.

Нейронная сеть, полученная в результате обучения, выражает закономерности, присутствующие в данных. Она оказывается функциональным эквивалентом некоторой модели зависимостей между переменными, подобной тем, которые строятся в традиционном моделировании. В отличие от традиционных моделей, выявленные зависимости не могут быть записаны в явном виде. В настоящее время существует достаточно много эффективных и разнообразных методов создания нейросетевой модели [4-8].

3. Data Mining в реальном времени

Как известно, статическими методами Data Mining выполняется анализ исторических данных, т. е. данных, накопленных ранее и неизменяемых в про-

цессе анализа. Методы Data Mining в реальном времени (Real-Time Analytics), в отличие от статических методов обучаются динамически и основаны на обратной связи от прогноза, полученного с помощью предсказательной модели. Этот подход поддерживается постоянным переобучением модели. Чаще всего приложения реального времени используют новые данные [3].

Адаптивный подход к обучению модели заключается в том, что она динамически переобучивается с использованием новых данных, получаемых в режиме реального времени. Достоинства адаптивного подхода состоят в том, что требуется меньше статистических предположений, модели быстро приспосабливаются к изменяющейся окружающей среде, а системы могут систематически собирать и анализировать новые данные.

4. Гибридная модель анализа больших данных

В настоящей работе предлагается гибридная модель для интеллектуального анализа больших данных в реальном времени. Весь процесс анализа данных разбит на следующие этапы: кластерный анализ, построение нейронной сети, выделение обучающих выборок и инициализация нейронной сети, получение данных в реальном времени и подача их на вход нейронной сети, обработка результатов классификации.

Для кластеризации предлагается использование методов адаптивной кластеризации. Классификатор построен на основе нейронной сети. При этом решены задачи по сбору данных, характерных для предметной области и их разбивке на множества обучающих, контрольных и тестовых данных. Выполнена предварительная обработка данных, выбраны признаки для подачи на вход сети, характерные для данной задачи. Были подобраны топология сети и функция активации. Рассматривалась структура многослойной сети. Была организована фаза обучения сети на основе установленных параметров, разработаны и внедрены типичные примеры использования сети. Было оценено качество работы сети с целью оптимизации архитектуры сети. При этом использовалась система реальных тестовых данных для определения и оценки погрешности.

Для проведения сравнительного анализа были реализованы некоторые классические статистические модели [9, 10].

Вышеуказанные задачи реализованы в соответствующих блоках разработанной системы. В фазе обучения использовались обучающие последовательности из разных классов. Рассматривались регрессионные и авторегрессионные модели (ARIMA). Нейронные сети выбранной архитектуры использовались для прогнозирования температуры воздуха, индексов потребительских цен и курсов валют. При необходимости в системе одна и та же задача решалась разными методами.

Эксперименты показали, что в отличие от статистического анализа, основанного на общих рекомендациях, нейросети способны в большинстве случаев строить оптимальную модель прогнозирования, более того, модель адаптивна и меняется динамично вместе с данными, что особенно важно для применения [10].

Помимо повышения скорости анализа данных этот подход имеет и еще одно важное преимущество - несколько относительно простых моделей по отдельности легче создавать и поддерживать, чем одну большую. Можно запускать модели поэтапно, получая таким образом первые результаты в сжатые сроки.

Предлагаемый подход по созданию гибридной модели даст возможность классифицировать объект в реальном времени, относить его к некоторому кластеру, и поскольку каждый кластер имеет специализированное описание и методы анализа данных, то в реальном времени можно формировать соответствующие рекомендательные системы для новых данных.

Список литературы:

1. Арустамов А, Анализ больших объемов данных, BaseGroup Labs, 2016.
2. Елизаров С.И. Кластеризация данных при помощи нечетких отношений, СПб.: БХВ, Петербург, 2004.
3. J. Garcke, M. Griebel, and M. Thess. Data Mining with Sparse Grids. In Computing, 2001.
4. Боровиков В.П. Нейронные сети. Методология и технологии современного анализа данных, изд. – М. Горячая линия, Телеком, 2008.
5. Комашинский В.И. Нейронные сети и их применение в системах управления и связи. М, Горячая линия, Телеком 2003.

6. Zhang G., Patuwo B. E., Michael Y. H. Forecasting with artificial neural networks, International Journal of Forecasting. Vol. 14, Issue 1.1998.
7. Ежов А.А. Нейрокомпьютинг и его применения в экономике и бизнесе, М.: МИФИ, 1998.
8. Саркисян С.Г, Овакимян А.С, Хачоян Л.О. Применение нейронных сетей для анализа больших данных. Сборник материалов ЕМС-форума, г. Ялта, 2013г.
9. Овакимян А.С, Саркисян С.Г, Зироян М.А., Тинякова В.И., Проектирование систем принятия решений, функционирующих в условиях неопределенности. Современная экономика: проблемы и решения. г. Воронеж, 2014 г.,
10. Харатян А., Саркисян С.Г, Овакимян А.С, Прогнозирование временных рядов с использованием искусственных нейронных сетей, ЕГУ, факультет экономики. Ежегодник, Ереван, 2014.