

SPRINGER BRIEFS IN COMPUTER SCIENCE

Filippo Maria Bianchi

Enrico Maiorino

Michael C. Kampffmeyer

Antonello Rizzi

Robert Jensen

Recurrent Neural Networks for Short-Term Load Forecasting

An Overview and
Comparative Analysis



Springer

SpringerBriefs in Computer Science

Series editors

Stan Zdonik, Brown University, Providence, Rhode Island, USA

Shashi Shekhar, University of Minnesota, Minneapolis, Minnesota, USA

Xindong Wu, University of Vermont, Burlington, Vermont, USA

Lakhmi C. Jain, University of South Australia, Adelaide, South Australia, Australia

David Padua, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

Xuemin (Sherman) Shen, University of Waterloo, Waterloo, Ontario, Canada

Borko Furht, Florida Atlantic University, Boca Raton, Florida, USA

V.S. Subrahmanian, University of Maryland, College Park, Maryland, USA

Martial Hebert, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Katsushi Ikeuchi, University of Tokyo, Tokyo, Japan

Bruno Siciliano, Università di Napoli Federico II, Napoli, Italy

Sushil Jajodia, George Mason University, Fairfax, Virginia, USA

Newton Lee, Newton Lee Laboratories, LLC, Tujunga, California, USA

More information about this series at <http://www.springer.com/series/10028>

Filippo Maria Bianchi · Enrico Maiorino
Michael C. Kampffmeyer
Antonello Rizzi · Robert Jenssen

Recurrent Neural Networks for Short-Term Load Forecasting

An Overview and Comparative Analysis

Filippo Maria Bianchi
UiT The Arctic University of Norway
Tromsø
Norway

Antonello Rizzi
Sapienza University of Rome
Rome
Italy

Enrico Maiorino
Harvard Medical School
Boston, MA
USA

Robert Jenssen
UiT The Arctic University of Norway
Tromsø
Norway

Michael C. Kampffmeyer
UiT The Arctic University of Norway
Tromsø
Norway

ISSN 2191-5768 ISSN 2191-5776 (electronic)
SpringerBriefs in Computer Science
ISBN 978-3-319-70337-4 ISBN 978-3-319-70338-1 (eBook)
<https://doi.org/10.1007/978-3-319-70338-1>

Library of Congress Control Number: 2017957698

© The Author(s) 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The key component in forecasting demand and consumption of resources in a supply network is an accurate prediction of real-valued time series. Indeed, both service interruptions and resource waste can be reduced with the implementation of an effective forecasting system. Significant research has thus been devoted to the design and development of methodologies for short-term load forecasting over the past decades. A class of mathematical models, called recurrent neural networks, are nowadays gaining renewed interest among researchers and they are replacing many practical implementations of the forecasting systems, previously based mostly on statistical methods. Despite the undeniable expressive power of these architectures, their recurrent nature complicates their understanding and poses challenges in the training procedures. Although recently different kinds of recurrent neural networks have been successfully applied in fields like natural language processing or text translation, a systematic evaluation of their performance in the context of load forecasting is still lacking. In this work, we perform a comparative study on the problem of short-term load forecast, by using different classes of state-of-the-art recurrent neural networks. We provide a general overview of the most important architectures and we define guidelines for configuring the recurrent networks to predict real-valued time series. We test the reviewed models on controlled synthetic tasks and on real-world datasets, covering important practical case studies. It is our hope that this essay can become a useful resource for data scientists in academia and industry to keep up-to-date with the latest developments in the field of deep learning and time series prediction.

Tromsø, Norway
September 2017

Filippo Maria Bianchi

Contents

1	Introduction	1
	References	5
2	Properties and Training in Recurrent Neural Networks	9
2.1	Backpropagation Through Time	11
2.2	Gradient Descent and Loss Function	13
2.3	Parameters Update Strategies	14
2.4	Vanishing and Exploding Gradient	17
	References	19
3	Recurrent Neural Network Architectures	23
3.1	Elman Recurrent Neural Network	23
3.2	Long Short-Term Memory	25
3.3	Gated Recurrent Unit	26
	References	28
4	Other Recurrent Neural Networks Models	31
4.1	NARX Network	31
4.2	Echo State Network	34
	References	37
5	Synthetic Time Series	41
	References	42
6	Real-World Load Time Series	45
6.1	Orange Dataset—Telephonic Activity Load	45
6.2	ACEA Dataset—Electricity Load	48
6.3	GEFCom2012 Dataset—Electricity Load	52
	References	55

- 7 Experiments** 57
 - 7.1 Experimental Settings 58
 - 7.1.1 ERNN, LSTM, and GRU 58
 - 7.1.2 NARX 59
 - 7.1.3 ESN 60
 - 7.2 Results on Synthetic Dataset 60
 - 7.3 Results on Real-World Dataset 63
 - 7.3.1 Results on Orange Dataset 63
 - 7.3.2 Results on ACEA Dataset 65
 - 7.3.3 Results on GEFCom Dataset 67
 - References 68

- 8 Conclusions** 71
 - Reference 72