

FinNLP-Muffin 2023

**Joint Workshop of the 5th Financial Technology  
and Natural Language Processing (FinNLP) and  
2nd Multimodal AI For Financial Forecasting  
(Muffin) in conjunction with IJCAI 2023**

**Proceedings of the Workshop**

August 20, 2023

Macao



# Preface

Welcome to FinNLP, a forum dedicated to fostering international collaboration and knowledge-sharing in the application of Natural Language Processing (NLP) within the dynamic domain of FinTech. As we gather in this workshop, our aim is to explore the intersection of FinTech and NLP, identifying challenging problems, shaping future research directions, and expanding the horizons of this interdisciplinary field. The 5th FinNLP is a joint workshop with the 2nd workshop on Multimodal AI For Financial Forecasting (Muffin).

In the current year, FinNLP-Muffin continues to be enriched by the participants, who bring forth novel ideas and share their latest findings on pertinent issues. Once again, we are thrilled to collaborate with 3DS Outscale (formerly Fortia) in organizing the shared tasks in FinNLP, with a focus on Multi-Lingual ESG Issue Identification (ML-ESG).

This workshop would not have been possible without the contributions of numerous individuals, and we extend our heartfelt gratitude to each of them. Dr. Juyeon Kang, who led a dedicated team in successfully orchestrating the shared task. Additionally, we are indebted to all the program committee members who invested substantial time and expertise in providing insightful comments on the submissions and guiding the selection process for FinNLP-Muffin-2023.

Last but not least, we extend our sincere thanks to the project JPNP20006, commissioned by the New Energy and Industrial Technology Development Organization (NEDO), for their invaluable financial support. Their partnership has been instrumental in enabling us to realize the goals of FinNLP-Muffin and advance research in this domain.

We hope that FinNLP-Muffin will continue to serve as a catalyst for groundbreaking research and meaningful connections, propelling the realms of FinTech and NLP towards new heights of innovation and excellence.

Chung-Chi Chen, Puneet Mathur, Ramit Sawhney, Hiroya Takamura  
FinNLP-Muffin 2023 Organizers  
August 2023



# Organizations

## **General Chairs, FinNLP**

Chung-Chi Chen, AIST, Japan

Hiroya Takamura, AIST, Japan

## **General Chairs, Muffin**

Puneet Mathur, University of Maryland College Park, USA

Ramit Sawhney, Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi

## **Organizing Committee**

Dinesh Manocha, University of Maryland College Park, USA

Preslav Nakov, Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi

Hen-Hsen Huang, Institute of Information Science, Academia Sinica, Taiwan

Hsin-Hsi Chen, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan

Hiroki Sakaji, School of Engineering, The University of Tokyo, Japan

Kiyoshi Izumi, School of Engineering, The University of Tokyo, Japan

## **Advisory Committee**

Franck Dernoncourt, Adobe Research, USA

Fu-Ming Guo, Fidelity Investments, USA

Lucie Flek, University of Marburg, Germany

Sanghamitra Dutta, University of Maryland College Park, USA

Sudheer Chava, Georgia Institute of Technology, USA

## **Shared Task Organizers**

Chung-Chi Chen, AIST, Japan

Yu-Min Tseng, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan

Juyeon Kang, 3DS Outscale (ex Fortia)

Anaïs Lhuissier, 3DS Outscale (ex Fortia)

Min-Yuh Day, Graduate Institute of Information Management, National Taipei University

Teng-Tsai Tu, Graduate Institute of International Business, National Taipei University

Hsin-Hsi Chen, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan

**Program Committee**

Paulo Alves, Católica Porto Business School

K. Selcuk Candan, Arizona State University

Damir Cavar, Indiana University Bloomington

Yung-Chun Chang, Taipei Medical University

Emmanuele Chersoni, Aix-Marseille University

Nelson Correa, Andinum, Inc.

Pablo Duboue, Textualization Software Ltd.

Jinhang Jiang, Walmart Inc.

Chit-Kwan Lin, Revelata, Inc.

Chenyang Lyu, Dublin City University

Hilal Pataci, Rensselaer Polytechnic Institute

Yulong Pei, Eindhoven University of Technology

Ke Tian, Rakuten Group, Inc

Chuan-Ju Wang, Academia Sinica

Shih-Hung Wu, Dept. of CSIE, Chaoyang University of Technology

Annie T.T. Ying, IBM

# Table of Contents

## **Model-Agnostic Meta-Learning for Natural Language Understanding Tasks in Finance**

Bixing Yan, Shaoling Chen, Yuxuan He and Zhihan Li ----- 1

## **ChatGPT as Data Augmentation for Compositional Generalization: A Case Study in Open Intent Detection**

Yihao Fang, Xianzhi Li, Stephen Thomas and Xiaodan Zhu ----- 13

## **Beyond Classification: Financial Reasoning in State-of-the-Art Language Models**

Guijin Son, Hanearl Jung, Moonjeong Hahm, Keonju Na and Sol Jin ----- 34

## **Textual Evidence Extraction for ESG Scores**

Naoki Kannan and Yohei Seki ----- 45

## **A Scalable and Adaptive System to Infer the Industry Sectors of Companies: Prompt + Model Tuning of Generative Language Models**

Lele Cao, Vilhelm von Ehrenheim, Astrid Berghult, Cecilia Henje, Richard Anselmo Stahl, Joar Wandborg, Sebastian Stan, Armin Catovic, Erik Ferm and Hannes Ingelhart ----- 55

## **Using Deep Learning to Find the Next Unicorn: A Practical Synthesis on Optimization Target, Feature Selection, Data Split and Evaluation Strategy**

Lele Cao, Vilhelm von Ehrenheim, Sebastian Krakowski, Xiaoxue Li and Alexandra Lutz ----- 63

## **Breaking the Bank with ChatGPT: Few-Shot Text Classification for Finance**

Lefteris Loukas, Ilias Stogiannidis, Prodromos Malakasiotis and Stavros Vassos ----- 74

## **DeRisk: An Effective Deep Learning Framework for Credit Risk Prediction over Real-World Financial Data**

Yancheng Liang, Jiajie Zhang, Hui Li, Xiaochen Liu, Yi Hu, Yong Wu, Jiaoyao Zhang, Yongyan Liu and Yi Wu ----- 81

## **Reducing tokenizer's tokens per word ratio in Financial domain with T-MuFin BERT Tokenizer**

Braulio Blanco Lambruschini, Patricia Becerra-Sanchez, Mats Brorsson and Maciej Zurad ----- 94

## **LoKI: Money Laundering Report Generation via Logical Table-to-Text using Meta Learning**

Harika Cm, Debasmita Das, Ram Ganesh V, Rajesh Kumar Ranjan and Siddhartha Asthana ----- 104

## **Multi-Lingual ESG Issue Identification**

Chung-Chi Chen, Yu-Min Tseng, Juyeon Kang, Anaïs Lhuissier, Min-Yuh Day, Teng-Tsai Tu, Hsin-Hsi Chen ----- 111

## **Leveraging Contrastive Learning with BERT for ESG Issue Identification**

Weiwei Wang, Wenyang Wei, Qingyuan Song and Yansong Wang ----- 116

## **Leveraging BERT Language Models for Multi-Lingual ESG Issue Identification**

Elvys Linhares Pontes, Mohamed Benjannet and Lam Kim Ming ----- 121

**EaSyGuide : ESG Issue Identification Framework leveraging Abilities of Generative Large Language Models**

Hanwool Lee, Jonghyun Choi, Sohyeon Kwon and Sungbum Jung ----- 127

**Jetsons at the FinNLP-2023: Using Synthetic Data and Transfer Learning for Multilingual ESG Issue Classification**

Parker Glenn, Alolika Gon, Nikhil Kohli, Sihan Zha, Parag Pravin Dakle and Preethi Raghavan ----- 133

**HKESG at the ML-ESG Task: Exploring Transformer Representations for Multilingual ESG Issue Identification**

Ivan Mashkin and Emmanuele Chersoni ----- 140

**Team HHU at the FinNLP-2023 ML-ESG Task: A Multi-Model Approach to ESG-Key-Issue Classification**

Fabian Billert and Stefan Conrad ----- 146