MLiSE 2021

1st Workshop on Machine Learning in Software Engineering

in conjunction with ECML PKDD

17th September 2021

Motivation

- **Software engineering (SE)** is about methodologies and techniques for building high quality software systems.

- Modern software systems are becoming larger and more complex.

- Need to improve quality, reliability, cost-effectiveness and the ability to solve complex problems.

- ML has revolutionized numerous domains.

Motivation

- **ML in SE:**
  - source code generation from requirements, automatically proving the correctness of software specifications,
  - providing intelligent assistance to developers.

- **SE in ML**
  - SE techniques and methodologies can be used to improve the ML process

  Goal of this workshop: create a new forum to bring together the ML and SE communities

Organization committee

Philippe Fournier-Viger
Shenzhen University

M. Saqib Nawaz
Harbin Inst. of Techology

Sebastian Ventura
University of Cordoba

Meng Sun
Peking University

Min Zhou
Huawei Noah's Ark Lab

Program committee

- Moulay Akhloufi (University of Moncton, Canada)

- Guangdong Bai (University of Queensland, Australia)

- Mustapha Bouakkaz (Université de Laghouat, Algeria)

- Tin Truong Chi (University of Dalat, Vietnam)

- Wensheng Gan (Jinan University, China)

- Osman Hasan (National University of Science and Technology, Pakistan)

- Tzung-Pei Hong (National University of Kaohsiung, Taiwan)

- Rage Uday Kiran (University of Aizu, Japan)

- M. Ikram Ullah Lali (University of Education, Pakistan)

- Pinar Karagoz (Middle East Technical University, Turkey)

- Bac Le (Ho Chi Minh City University of Science, Vietnam)

- Jerry Chun-Wei Lin (Western Norway University of Applied Sciences, Norway)

- Ai Liu (Hiroshima University, Japan)

- Wanwei Liu (NUDT, China)

- Jose Maria Luna (University of Cordoba, Spain)

- João Mendes-Moreira (University of Porto, Portugal)

- Engelbert Mephu Nguifo (Université Blaise Pascal, France)

- Farid Nouioua (University of Bordj Bou Arreridj, Algeria)

- José Proença (ISEP, Portugal)

- Wei Song (North China University of Technology)

- Bay Vo (Ho Chi Minch City University of Technology, Vietnam)

- Cheng-Wei Wu, (National Ilan University, Taiwan)

- Youxi Wu (Hebei University of Technology, China)

- Unil Yun (Sejong University, Korea)

- Xiyue Zhang (Peking University, China)

Program

- 2 keynote speakers

- 13 submitted papers

- Double-blind review

- 7 accepted papers (53%)

- The proceedings of **MLiSE** will be jointly published with other workshops after the conference, by Springer.

- https://philippe-fournier-viger.com/MLISE_workshop_2021/MLISE_proceedings.pdf

Keynote speaker 1:

Prof. Zhi Jin
(Peking University, China)

"Deep learning enabled program understanding"

Keynote speaker 2:

13:20 to 14:30

Prof. Atif Mashkoor
(Johannes Kepler University, Austria)

"The Relationship Between Machine Learning and Software Engineering Life Cycle Stages"

★ ★ ★

Best paper award

THIS ACKNOWLEDGES THAT THE PAPER

A Stacked Bidirectional LSTM Model for Classifying Source Codes Built in MPLs
by
Md. Mostafizer Rahman, Yutaka Watanobe, Rage Uday Kiran and Raihan Kabir

HAS BEEN SELECTED AS THE BEST PAPER AT THE 1ST WORKSHOP ON MACHINE
LEARNING IN SOFTWARE ENGINEERING (MLISE 2021)
HELD ONLINE AT THE ECML PKDD 2021 CONFERENCE

Philippe

_____

Philippe Fournier-Viger
on behalf of the MLiSE 2021 committee

17 September 2021

_____

Date

Research papers

10:30-11:00	**Paper:** A Stacked Bidirectional LSTM Model for Classifying Source Codes Built in MPLs. (best paper)	Md. Mostafizer Rahman, Yutaka Watanobe, Rage Uday Kiran and Raihan Kabir.
11:00-11:30	**Paper**: Metamorphic Malware Behavior Analysis using Sequential Pattern Mining	M. Saqib Nawaz, Philippe Fournier-Viger, M. Zohaib Nawaz, Chen Guoting and Youxi Wu
11:30-12:00	**Paper**: Applying Machine Learning to Risk Assessment in Software Projects	André Sousa, João Faria, João Moreira, Duarte Gomes, Pedro Henriques and Ricardo Graça.

14:30-15:00	**Paper**: SampleFix: Learning to Generate Functionally Diverse Fixes	Hossein Hajipour, Apratim Bhattacharya, Cristian-Alexandru Staicu and Mario Fritz.
15:00-15:30	**Paper**: Linguistic Analysis of Stack Overflow Data: Native English vs Non-Native English Speakers	Janneke Morin and Kris Gosh
15:30-16:00	**Paper**: IReEn: Reverse-Engineering of Black-Box Functions via Iterative Neural Program Synthesis.	Hossein Hajipour, Mateusz Malinowski and Mario Fritz.
16:00-16:30	**Paper**: Machine Learning for Intelligent Industrial Design.	Philippe Fournier-Viger, M. Saqib Nawaz, Wei Song and Wensheng Gan

Next year

- Try to expand this workshop to a larger audience.
- A potential journal special issue