R2LM 2025

Proceedings of the First Workshop on Comparative Performance Evaluation: From Rules to Language Models

associated with

The 15th International Conference on Recent Advances in Natural Language Processing RANLP'2025

Edited by Alicia Picazo-Izquierdo, Ernesto Luis Estevanell-Valladares, Ruslan Mitkov, Rafael Munoz Guillena and Raul Garcia Cerda

11 September, 2025
Varna, Bulgaria

First Workshop on Comparative Performance Evaluation:
From Rules to Language Models
Associated with the International Conference
Recent Advances in Natural Language Processing
RANLP'2025

PROCEEDINGS

Varna, Bulgaria
11 September 2025

Online ISBN 978-954-452-102-8

Designed by INCOMA Ltd.
Shoumen, BULGARIA

Preface

The rapid growth of deep learning and large language models has transformed the field of natural language processing (NLP), driving remarkable progress in tasks that once seemed out of reach. Yet, as these models scale in size and capability, critical challenges remain. Issues of interpretability, robustness, long-context reasoning, and the substantial data and computational resources required continue to raise questions about their universality. At the same time, rule-based and knowledge-based approaches, once considered traditional or even outdated, are being reconsidered. Their strengths in precision, explainability, and adaptability to low-resource or domain-specific contexts make them valuable complements to data-driven methods. This workshop aims to unite researchers from various fields of symbolic, statistical, and hybrid methods to assess their comparative performance, understand their limitations, and explore how they can work together most effectively. By encouraging dialogue between different research methodologies and highlighting emerging trends such as retrieval-augmented generation, neurosymbolic AI, and knowledge graph-driven systems, we aim to establish a path towards more robust, efficient, and transparent NLP technologies. We hope that the contributions gathered here will not only advance methodological understanding but also inspire a more balanced and inclusive vision for the future of language technology.

Workshop Chairs

- Alicia Picazo-Izquierdo (University of Alicante)
- Ernesto Luis Estevanell-Valladares (University of Alicante, University of Havana)
- Ruslan Mitkov (Lancaster University, University of Alicante)
- Rafael Munoz Guillena (University of Alicante)
- Raul Garcia Cerda (University of Alicante)

Programme Committee

- Constantin Orasan (University of Surrey)
- Paul Greaney (Atlantic Technological University)
- Robiert Sepulveda (University of Alicante)
- Sandra Kuebler (Indiana University)
- Pablo Gervas (Complutense University of Madrid)
- Antonio Toral (University of Alicante)
- Aleksei Dorkin (University of Tartu)
- Sina Ahmadi (University of Zurich)
- Heili Orav (University of Tartu)
- Judith Sieker (Bielefeld University)

Table of Contents

CoVeGAT: A Hybrid LLM & Graph-Attention Pipeline for Accurate Citation-Aligned Claim Verification
Max Bader, Akshatha Arunkumar, Ohan Ahmad, Maruf Hassen, Charles Duong, Vasu Sharma, Sean O'Brien and Kevin Zhu 1

A Comparative Study of Vision Transformers and Multimodal Language Models for Violence Detection in Videos
Tomas Ditchfield-Ogle and Ruslan Mitkov 10

Detection of AI-generated Content in Scientific Abstracts
Ernesto Luis Estevanell-Valladares, Alicia Picazo-Izquierdo and Ruslan Mitkov 21

A Comparative Study of Hyperbole Detection Methods: From Rule-Based Approaches through Deep Learning Models to Large Language Models
Silvia Gargova, Nevena Grigorova and Ruslan Mitkov 30

Evaluating the Performance of Transformers in Translating Low-Resource Languages through Akkadian
Daniel A. Jones and Ruslan Mitkov 39

United We Fine-Tune: Structurally Complementary Datasets for Hope Speech Detection
Priya Dharshini Krishnaraj, Tulio Ferreira Leite da Silva, Gonzalo Freijedo Aduna, Samuel Chen, Farah Benamara and Alda Mari 48

Does Anaphora Resolution Improve LLM Fine-Tuning for Summarisation?
Yi Chun Lo and Ruslan Mitkov 59

Transformers and Large Language Models for Hope Speech Detection A Multilingual Approach
Diana Patricia Madera-Espíndola, Zoe Caballero-Domínguez, Valeria J. Ramírez-Macías, Sabur Butt and Hector G. Ceballos 67

Beyond BLEU: Ethical Risks of Misleading Evaluation in Domain-Specific QA with LLMs
Ayoub Nainia, Régine Vignes-Lebbe, Hajar Mousannif and Jihad Zahir 77

From Zero to Hero: Building Serbian NER from Rules to LLMs
Milica Ikonić Nešić, Sasa Petalinkar, Ranka Stanković and Ruslan Mitkov 87

Enhancing the Performance of Spoiler Review Detection by a LLM with Hints
Genta Nishi and Einoshin Suzuki 97

Evaluating Structured Decoding for Text-to-Table Generation: Evidence from Three Datasets
Julian Oestreich and Lydia Müller 113

Evaluating the LLM and NMT Models in Translating Low-Resourced Languages
Julita JP Pucinskaite and Ruslan Mitkov 123

KGEIR: Knowledge Graph-Enhanced Iterative Reasoning for Multi-Hop Question Answering
Tianda Sun and Dimitar Kazakov 134

From Handcrafted Features to LLMs: A Comparative Study in Native Language Identification
Aliyah C. Vanterpool and Katsiaryna Aharodnik 144

Systematic Evaluation of Rule-Based Analytics for LLM-Driven Graph Data Modelling
Fabio Antonio Yanez, Andrés Montoyo, Armando Suárez, Alejandro Piad-Morffis and Yudivián Almeida Cruz 154

Improved Contrastive Learning over Commonsense Knowledge Graphs for Unsupervised Reasoning
Rongwen Zhao and Jeffrey Flanigan 165

Conference Program

CoVeGAT: A Hybrid LLM & Graph-Attention Pipeline for Accurate Citation-Aligned Claim Verification
Max Bader, Akshatha Arunkumar, Ohan Ahmad, Maruf Hassen, Charles Duong, Vasu Sharma, Sean O'Brien and Kevin Zhu

A Comparative Study of Vision Transformers and Multimodal Language Models for Violence Detection in Videos
Tomas Ditchfield-Ogle and Ruslan Mitkov

16:30–17:00 *Detection of AI-generated Content in Scientific Abstracts*
Ernesto Luis Estevanell-Valladares, Alicia Picazo-Izquierdo and Ruslan Mitkov

11:00–11:30 *A Comparative Study of Hyperbole Detection Methods: From Rule-Based Approaches through Deep Learning Models to Large Language Models*
Silvia Gargova, Nevena Grigorova and Ruslan Mitkov

11:30–12:00 *Evaluating the Performance of Transformers in Translating Low-Resource Languages through Akkadian*
Daniel A. Jones and Ruslan Mitkov

United We Fine-Tune: Structurally Complementary Datasets for Hope Speech Detection
Priya Dharshini Krishnaraj, Tulio Ferreira Leite da Silva, Gonzalo Freijedo Aduna, Samuel Chen, Farah Benamara and Alda Mari

Does Anaphora Resolution Improve LLM Fine-Tuning for Summarisation?
Yi Chun Lo and Ruslan Mitkov

Transformers and Large Language Models for Hope Speech Detection A Multilingual Approach
Diana Patricia Madera-Espíndola, Zoe Caballero-Domínguez, Valeria J. Ramírez-Macías, Sabur Butt and Hector G. Ceballos

12:00–12:30 *Beyond BLEU: Ethical Risks of Misleading Evaluation in Domain-Specific QA with LLMs*
Ayoub Nainia, Régine Vignes-Lebbe, Hajar Mousannif and Jihad Zahir

From Zero to Hero: Building Serbian NER from Rules to LLMs
Milica Ikonić Nešić, Sasa Petalinkar, Ranka Stanković and Ruslan Mitkov

Enhancing the Performance of Spoiler Review Detection by a LLM with Hints
Genta Nishi and Einoshin Suzuki

Evaluating Structured Decoding for Text-to-Table Generation: Evidence from Three Datasets
Julian Oestreich and Lydia Müller

17:00–17:30 *Evaluating the LLM and NMT Models in Translating Low-Resourced Languages*
Julita JP Pucinskaite and Ruslan Mitkov

15:00–15:30 *KGEIR: Knowledge Graph-Enhanced Iterative Reasoning for Multi-Hop Question Answering*
Tianda Sun and Dimitar Kazakov

17:30–18:00 *From Handcrafted Features to LLMs: A Comparative Study in Native Language Identification*
Aliyah C. Vanterpool and Katsiaryna Aharodnik

15:30–16:00 *Systematic Evaluation of Rule-Based Analytics for LLM-Driven Graph Data Modelling*
Fabio Antonio Yanez, Andrés Montoyo, Armando Suárez, Alejandro Piad-Morffis and Yudivián Almeida Cruz

Improved Contrastive Learning over Commonsense Knowledge Graphs for Unsupervised Reasoning
Rongwen Zhao and Jeffrey Flanigan

TextBandit: Evaluating Probabilistic Reasoning in LLMs Through Language-Only Decision Tasks
Arjun Damerla, Jimin Lim, Yanxi Jiang, Nam Nguyen Hoai Le and Nikil Selladurai