

Title: Learning Representations of Learning Representations

Authors: Rita González-Márquez, Dmitry Kobak

Abstract: The ICLR conference is unique among the top machine learning conferences in that all submitted papers are openly available. Here we present the ICLR dataset consisting of abstracts of all ~24 thousand ICLR submissions from 2017-2024 with meta-data, decision scores, and custom keyword-based labels. We find that on this dataset, bag-of-words representation outperforms most dedicated sentence transformer models.