

# Explainable AI for Source Code Applications

...

sdmay25-30

# Overview

- Client/Company: Dr. Ali Jannesari/ISU SWAPP Lab
- Abstract:
  - Focus on auto-labeling code datasets using AST tools, regular expressions, and LLM-generated labels.
- Goal:
  - Enhance model interpretability by evaluating learned concepts against human-defined code properties.

# Problem Statement & Related Products

- Problem:
  - Current AI models for source code lack clear interpretability, hindering trust and understanding.
- Related Products:
  - ConceptX:
    - Analyzes encoded concepts in transformer language models.
  - NeuroX:
    - Provides neuron-level interpretation of NLP models.
- Product Limitations:
  - Both tools face high computational demands and have gaps in explaining complex, novel concept

# Market Gap & New Ideas

- Market Gap:
  - Lack of tools that integrate well with high-performance computing for large-scale source code interpretation.
  - Need for a more scalable, interpretable AI solution for code datasets.
- New Ideas from Research:
  - Improve scalability and user-friendliness for high-performance computing.
  - Create a more comprehensive web interface for visualizing learned concepts

# Conclusions

- Comparison:
  - Current solutions like ConceptX and NeuroX provide a solid foundation but lack in scalability and computational efficiency.
- Innovative Approach:
  - The proposed project aims to extend existing frameworks by addressing compatibility issues and enhancing usability for larger models and datasets