

Network Text Analysis Framework for Mapping Research Trends: A Neural Architecture Search Case Study

Petra Vidnerová, Roman Neruda
Institute of Computer Science, The Czech Academy of Sciences

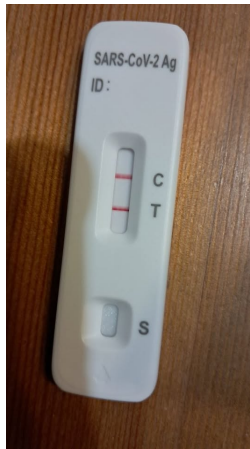
September 29, ITAT 2025



Spolufinancováno
Evropskou unií



Sorry for presenting online ...



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Outline

- ▶ (Network) text analysis
- ▶ Neural architecture search (NAS) research field
- ▶ Network text analysis applied to NAS
- ▶ Conclusion



Spolufinancováno
Evropskou unií

MS
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Text Analysis

Definition

Text Analysis is the process of automatically extracting meaningful information from unstructured text.

Key Goals

- ▶ Identify patterns, trends, and relationships in text.
- ▶ Transform unstructured data into structured insights.
- ▶ Support decision-making through language understanding.

Common Applications

- ▶ Sentiment analysis (e.g., customer feedback, social media).
- ▶ Topic modeling and classification.
- ▶ Information retrieval and summarization.
- ▶ Detecting relationships in large text corpora.



Spolufinancováno
Evropskou unií

MŠMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

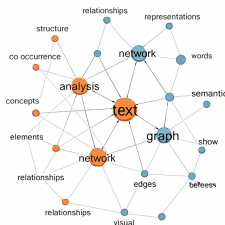
Network Text Analysis

Definition

Network Text Analysis combines text analysis with network theory to model texts as networks of concepts, terms, or entities.

Core Idea

- ▶ Words or concepts are represented as nodes.
- ▶ Their relationships (co-occurrence, similarity, or semantic links) form edges.
- ▶ The resulting network reveals structures of meaning within the text.



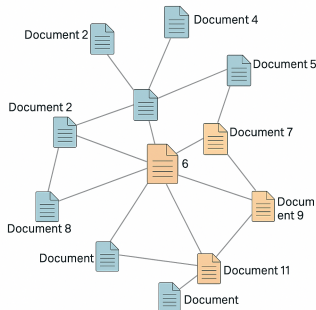
Spolufinancováno
Evropskou unií

MT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Network Text Analysis

Core Idea

- ▶ Documents are represented as nodes.
- ▶ Their relationships (based on similarity, shared keywords, citations, or thematic links) form edges.
- ▶ The resulting network reveals clusters of related texts and the overall structure of the knowledge domain.



Spolufinancováno
Evropskou unií

MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Network Text Analysis

Key Benefits

- ▶ Captures relational structures, not just word counts.
- ▶ Identifies central concepts and hidden connections.
- ▶ Helps map knowledge domains and discourse communities.

Applications

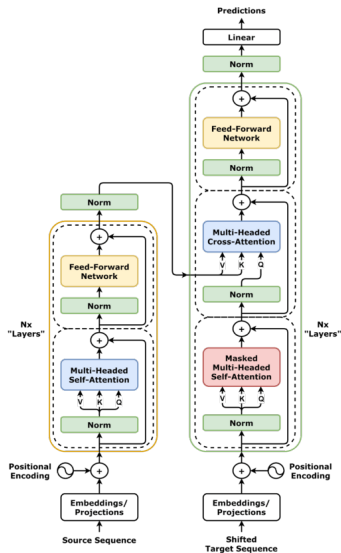
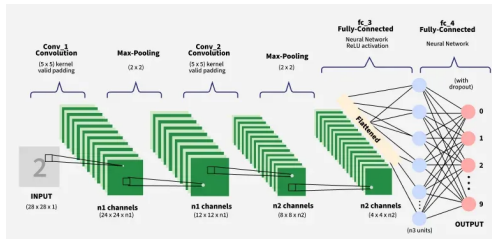
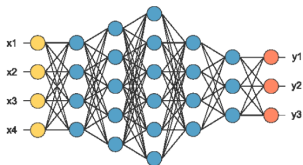
- ▶ Scientific literature mapping.
- ▶ Analyzing policy documents or news coverage.
- ▶ Exploring online discussions and social media.



Spolufinancováno
Evropskou unií


MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

(Deep) Neural Networks

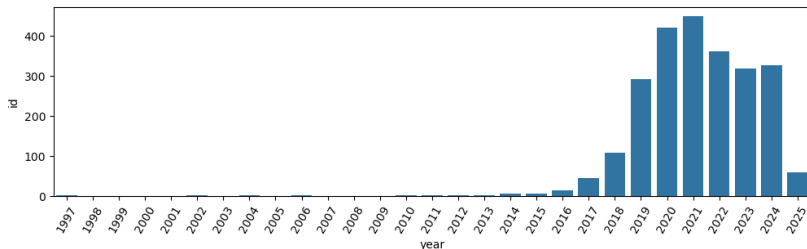


Neural Architecture Search

Neural Architecture Search (NAS)

Automating the design of neural network architecture. Given a problem, NAS looks for an optimal architecture.

Research in NAS



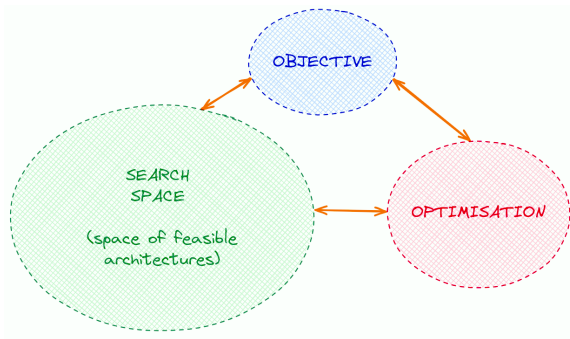
Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

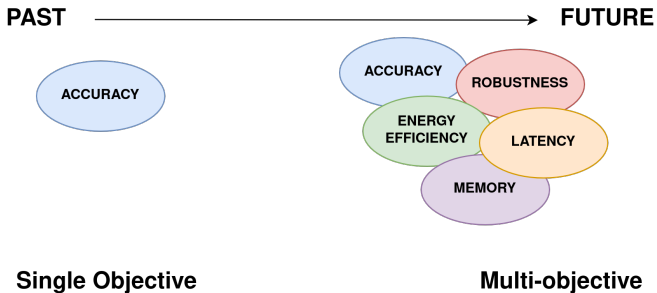
NAS – Optimisation Problem

- ▶ Minimize given objectives over the given search space



Objectives

- Measure the quality of a solution



Spolufinancováno
Evropskou unií

MT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Text Analysis

Dataset

- ▶ 10000 papers downloaded from ArXiv via ArXiv API .
 - ▶ Search query used: [Neural Architecture Search](#).
 - ▶ Still not all of them match the topic, filtered using LLM.
 - ▶ Final dataset: 2 423 papers
-
- ▶ Citation information can be found in OpenAlex bibliographic database (<https://openalex.org/>).
 - ▶ OpenAlex accessible via API, Python library [pyalex](#).
 - ▶ Each paper matched with its OpenAlex counterpart, citation information obtained.



Keywords

Keywords Extraction

- ▶ LLM used to extract keywords from abstracts.
- ▶ Fixed keywords - for a given keyword and abstract we ask the LLM if the keyword is relevant to the abstracts.
- ▶ LLM generated - for a given abstract we ask the LLM to generate up to 5 keywords.
- ▶ We used combination of both.

Most Frequent Keywords

efficiency, accuracy, computational cost, performance prediction, hyperparameter optimisation, model compression, multi-objective optimisation, evolutionary algorithms

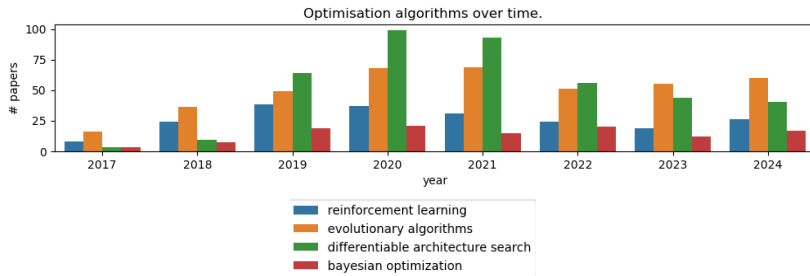


Spolufinancováno
Evropskou unií


MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

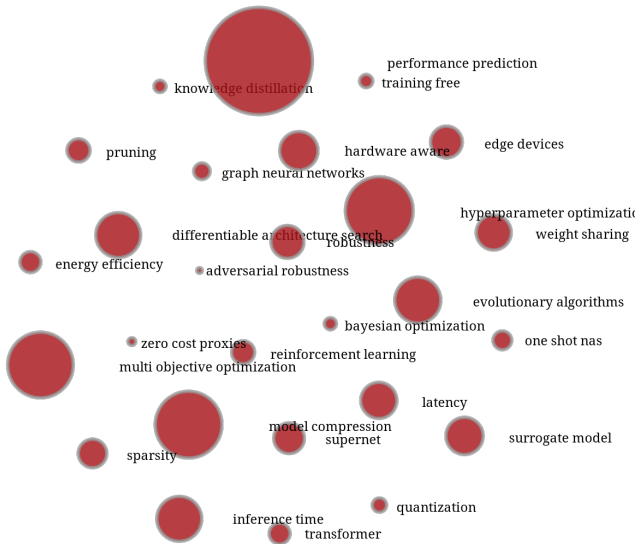
Optimisation

- ▶ Random search (baseline)
- ▶ Evolutionary and genetic algorithms
- ▶ Bayesian optimisation
- ▶ Reinforcement learning
- ▶ Differentiable techniques



Spolufinancováno
Evropskou unií

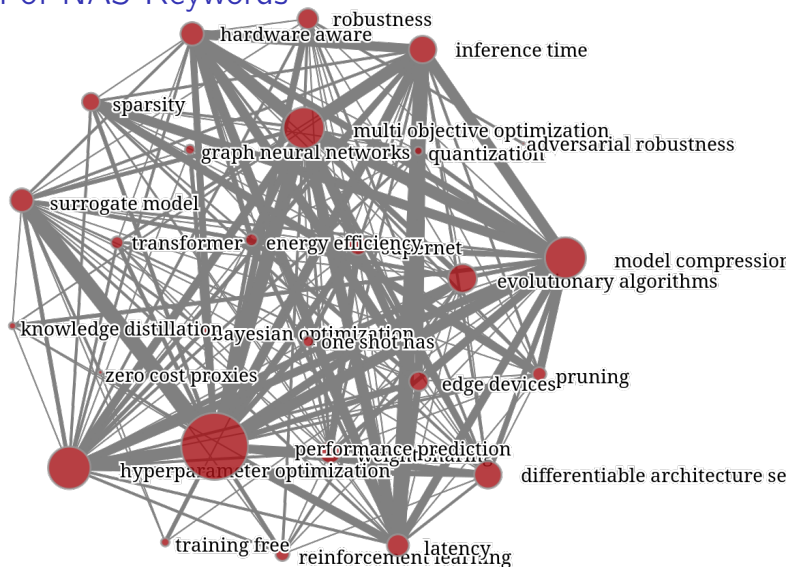
NAS Keywords



Spolufinancováno
Evropskou unií

MS
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Graph of NAS Keywords

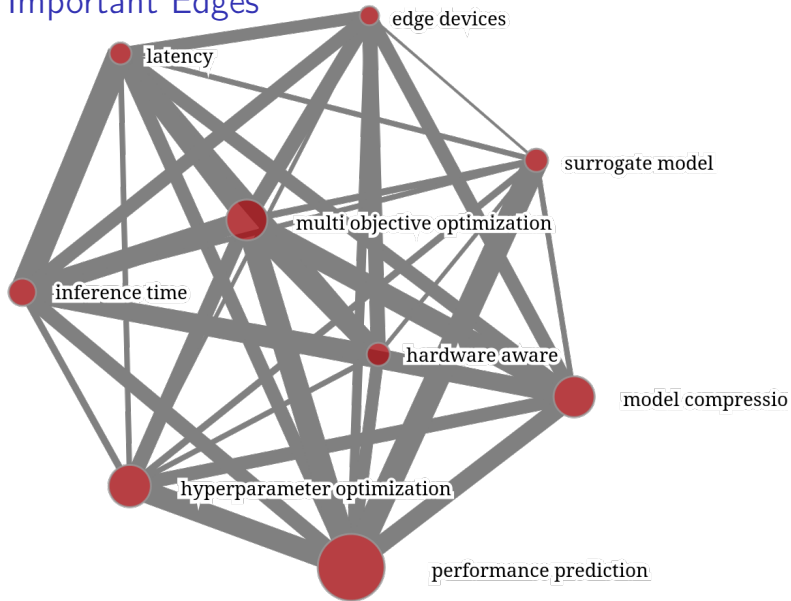


Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

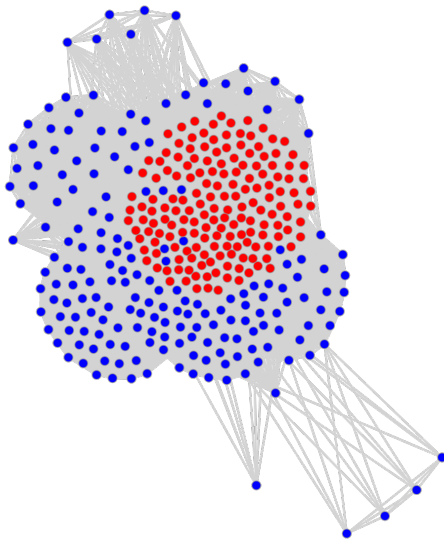
Most Important Edges



Spolufinancováno
Evropskou unií



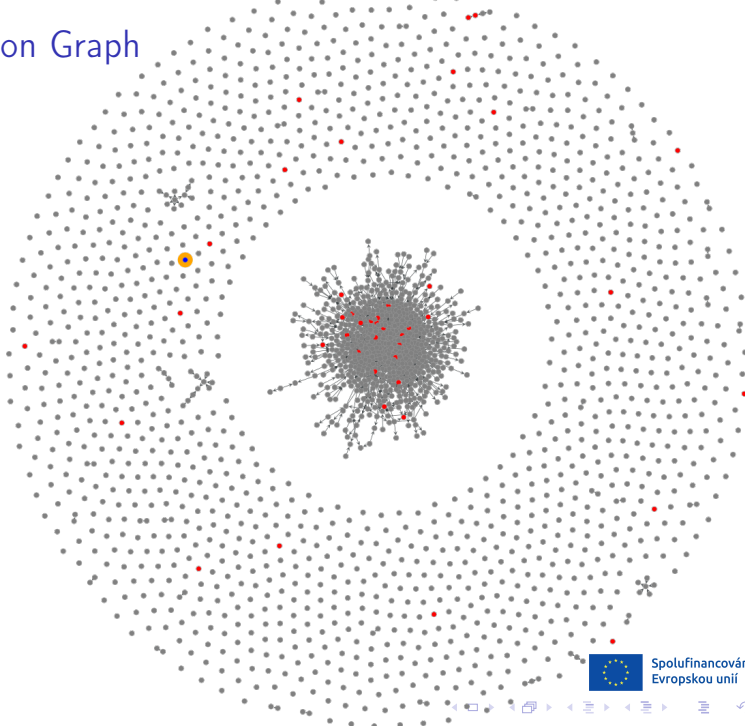
Paper Graph – Evolutionary Papers



Spolufinancováno
Evropskou unií

MT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

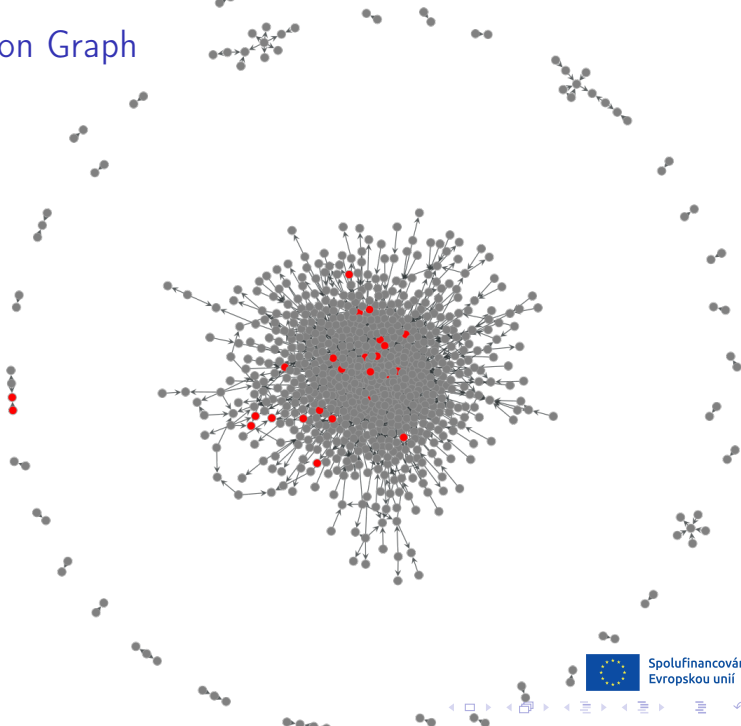
Citation Graph



Spolufinancováno
Evropskou unií

MŠMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

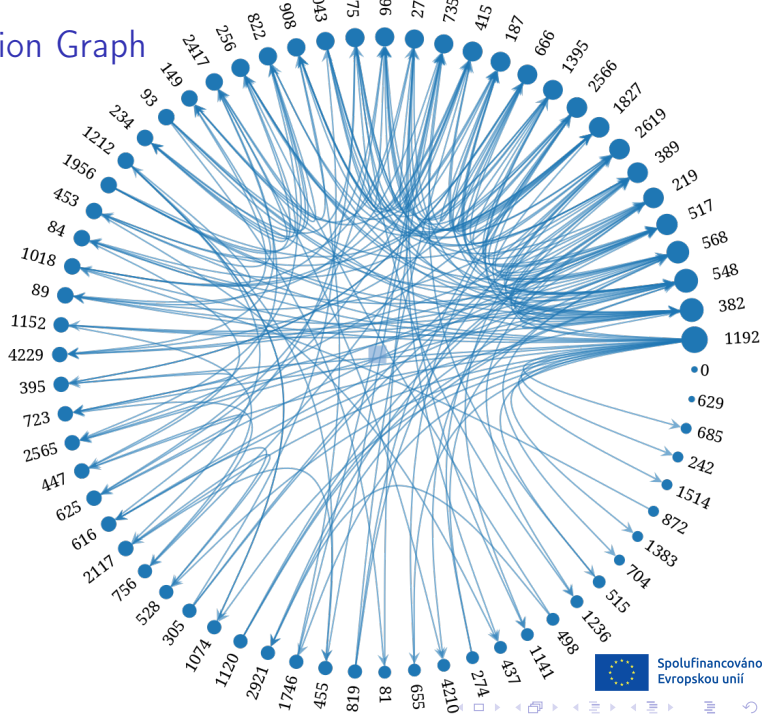
Citation Graph



Spolufinancováno
Evropskou unií

MŠMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Citation Graph



Spolufinancováno
Evropskou unií

Ministerstvo školství,
mládeže a tělovýchovy

Conclusion

- ▶ Methodology for research fields mapping proposed.
- ▶ Demonstrated on the NAS field.
- ▶ LLMs are helpful tools in text analysis.
- ▶ Networks help us to navigate through the field.
- ▶ Citation networks provide better insight than number of citation only, allow identification of key papers.



Spolufinancováno
Evropskou unií


MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Thank you! Questions?



Spolufinancováno
Evropskou unií

**MS
MT**
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



Spolufinancováno
Evropskou unií

**MS
MT**
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY