Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Leveraging Graph Neural Networks (GNNs) for Predicting Emerging Technologies

19. 06. 2025

Paul Bagourd, Julian Jang-Jaccard, Alain Mermoud (CYD Campus)

Research Motivations

Rapid publication growth

Scientific publications are growing at unprecedented rates.

Trend Tracking Challenge

Defense operations aim to identify relevant emerging trends.

Knowledge Graph

Existing methods like Knowledge Graphs map topic relationships.

GNN Opportunity

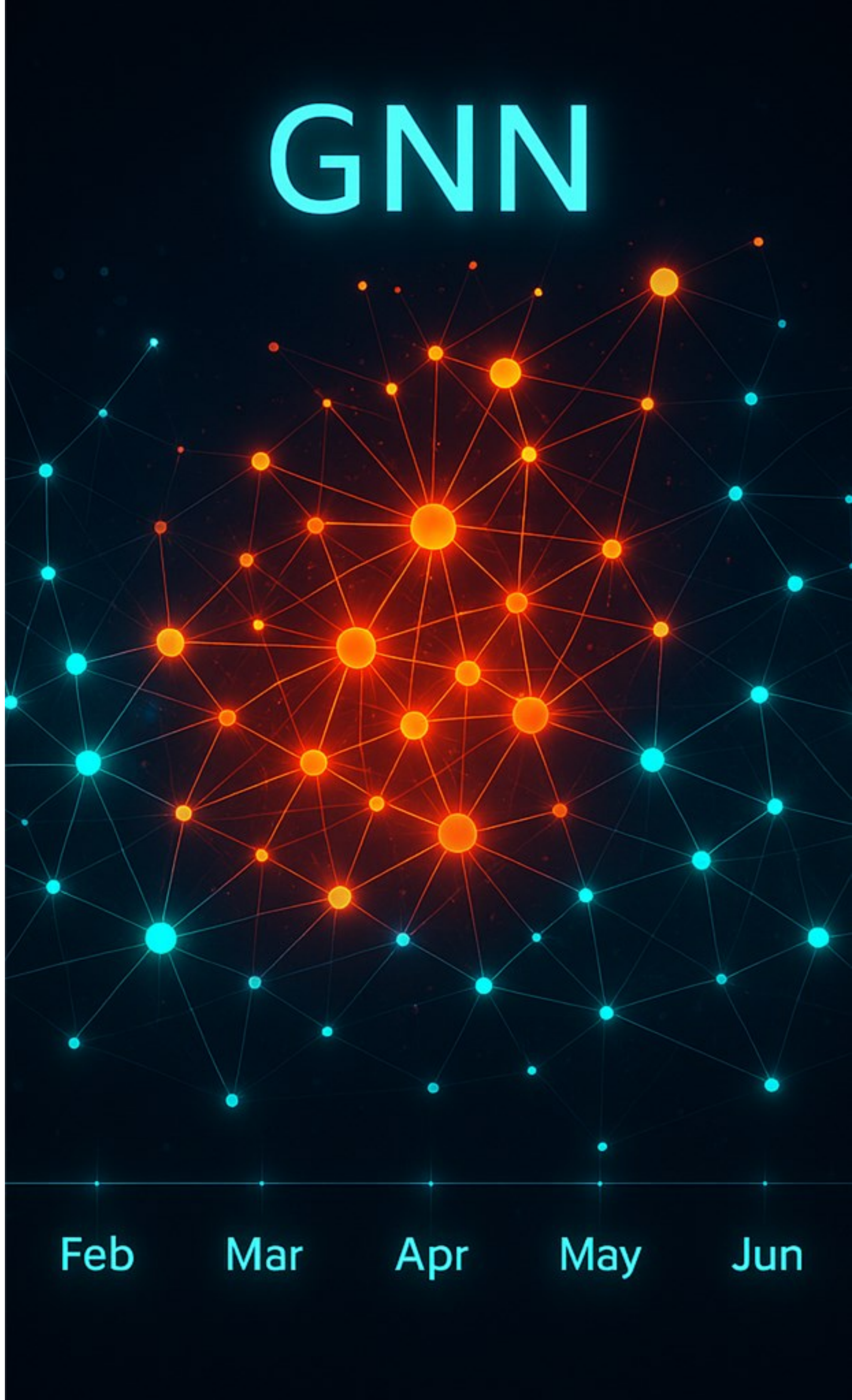
Can Graph Neural Networks (GNNs) enhance our ability to forecast the technological trends?

Research Objectives

Predict emerging technologies using GNNs

Leverage :

- ★ Frequency of specific trend mentions
- ★ Growth in citation rates
- ★ Long term dependencies across time



Data Collection and Processing

1. Data source

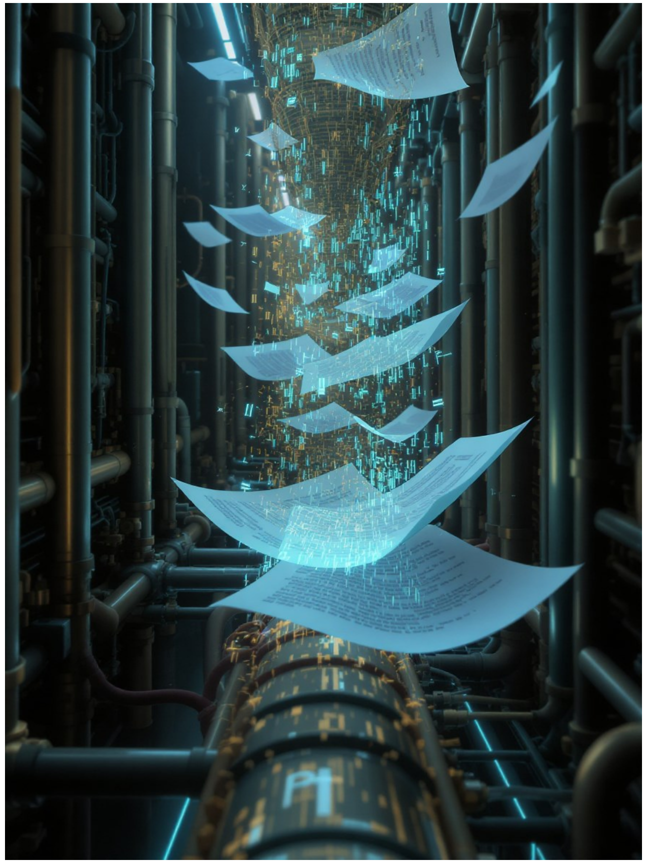
Abstracts of scientific paper from OpenAlex : select 100000 scientific papers (**quantum use case**)

2. Keyword extraction

Context aware extraction from the articles using NLP

KeyBERT : uses BERT embeddings, general keyword extraction
SPECTER : document-level embeddings for scientific papers

KeyBERT+SPECTER : extraction by leveraging the scientific context



Graph Construction

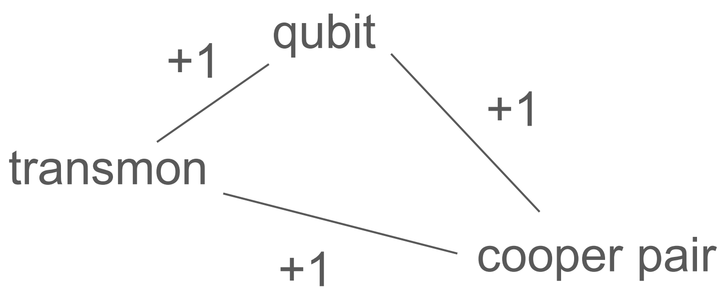
Charge insensitive qubit design derived from the Cooper pair box

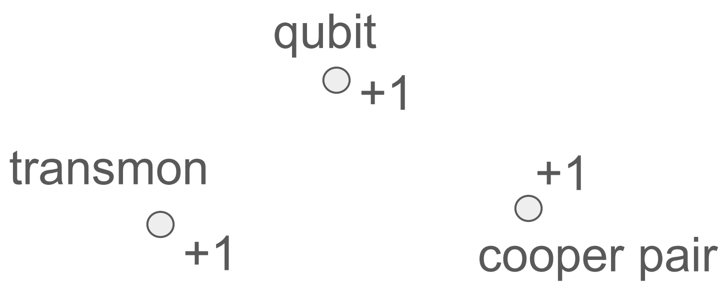
Jens Koch,[1] Terri M. Yu,[1] Jay Gambetta,[1] A. A. Houck,[1] D. I. Schuster,[1] J. Majer,[1] Alexandre Blais,[2] M. H. Devoret,[1] S. M. Girvin,[1] and R. J. Schoelkopf[1]

[1]Departments of Physics and Applied Physics, Yale University, New Haven, Connecticut 06520, USA
[2]Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1K 2R1
(Dated: September 26, 2007)

Short dephasing times pose one of the main challenges in realizing a quantum computer. Different approaches have been devised to cure this problem for superconducting qubits, a prime example being the operation of such devices at optimal working points, so-called "sweet spots." This latter approach led to significant improvement of T_2 times in Cooper pair box qubits [D. Vion et al., Science **296**, 886 (2002)]. Here, we introduce a new type of superconducting qubit called the "transmon." Unlike the charge qubit, the transmon is designed to operate in a regime of significantly increased ratio of Josephson energy and charging energy E_J/E_C. The transmon benefits from the fact that its charge dispersion decreases exponentially with E_J/E_C, while its loss in anharmonicity is described by a weak power law. As a result, we predict a drastic reduction in sensitivity to charge noise relative to the Cooper pair box and an increase in the qubit-photon coupling, while maintaining sufficient anharmonicity for selective qubit control. Our detailed analysis of the full system shows that this gain is not compromised by increased noise in other known channels.

PACS numbers: 03.67.Lx, 74.50.+r, 32.80.-t

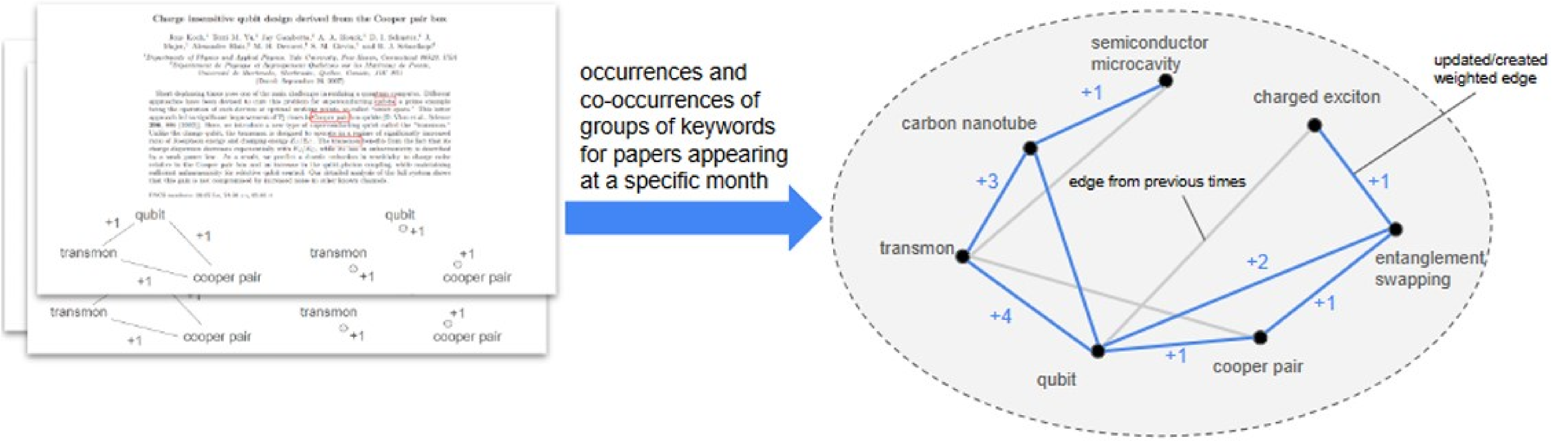




For each month :

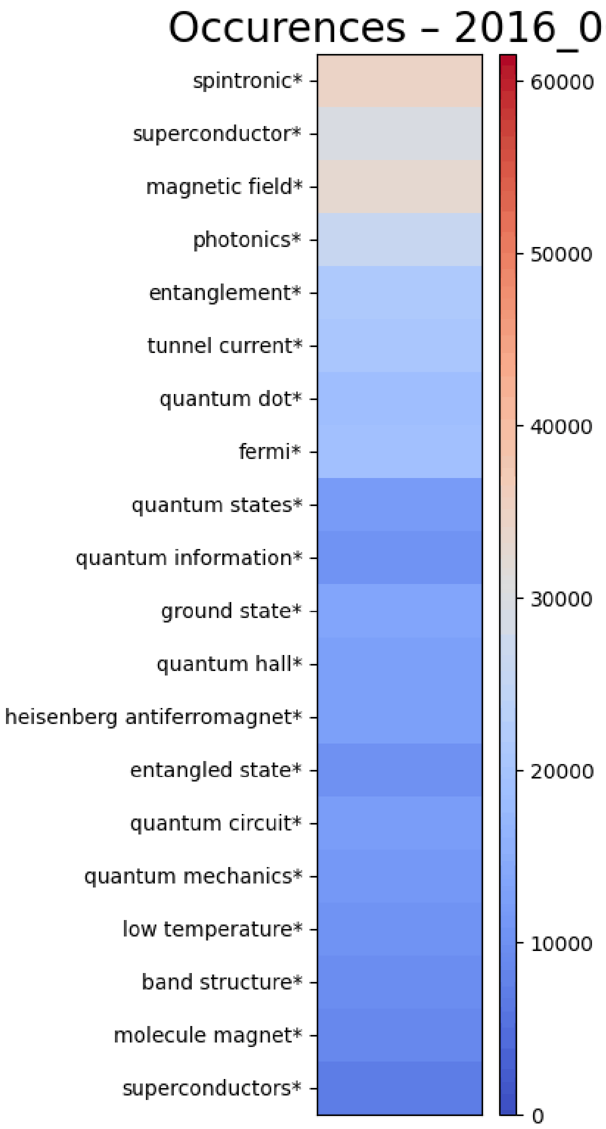
- Scan the abstracts of the new articles published
- Create a new graph

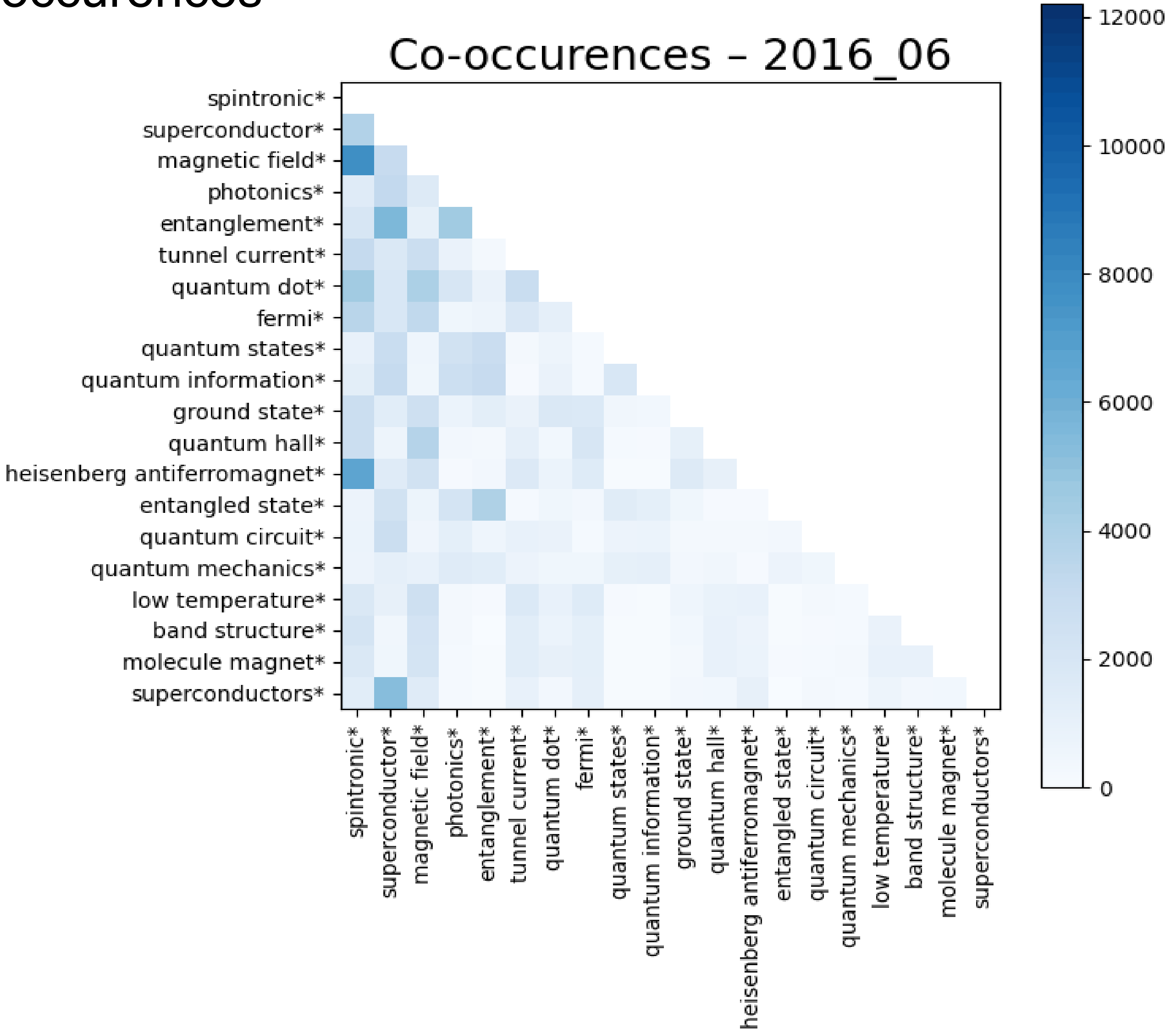
From 1987 to 2025 : 500 graphs



Identification Results

Cumulative Occurrences & Co-occurences





Prediction Task

Question : will a given node emerge in the future ?

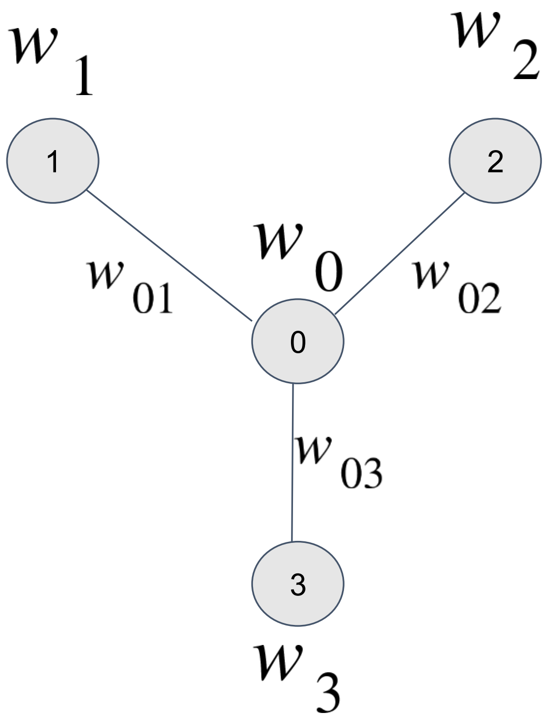
Definition of the emergence score

Percentage increase in occurences + co-occurences :

$$emergence(v, t+1) = \frac{w_v^{t+1}}{w_v^{t} + \varepsilon} + \frac{\sum_{u \in N(v)} w_{uv}^{t+1}}{\sum_{u \in N(v)} w_{uv}^{t} + \varepsilon}$$

Definition of emergence

Node emergence if score > 70th percentile (relative threshold)

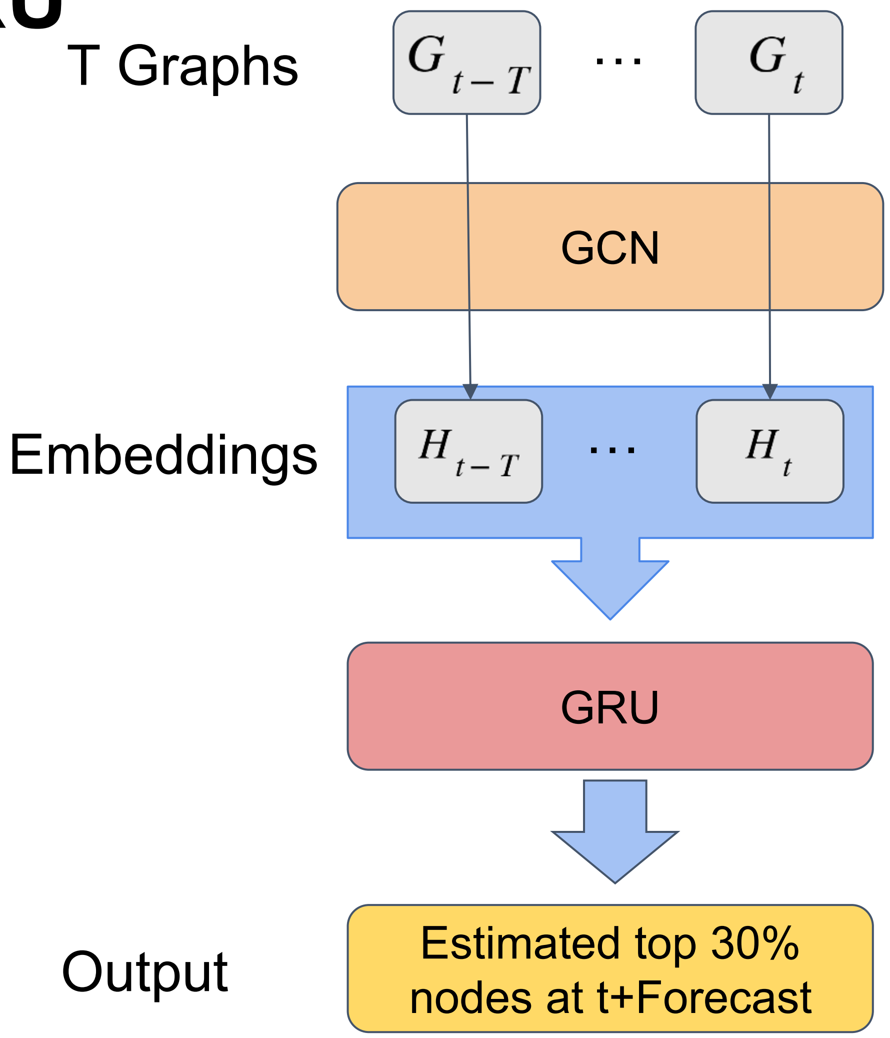


Our GNN model : GCN+GRU

Classification task: predict if a node will emerge.

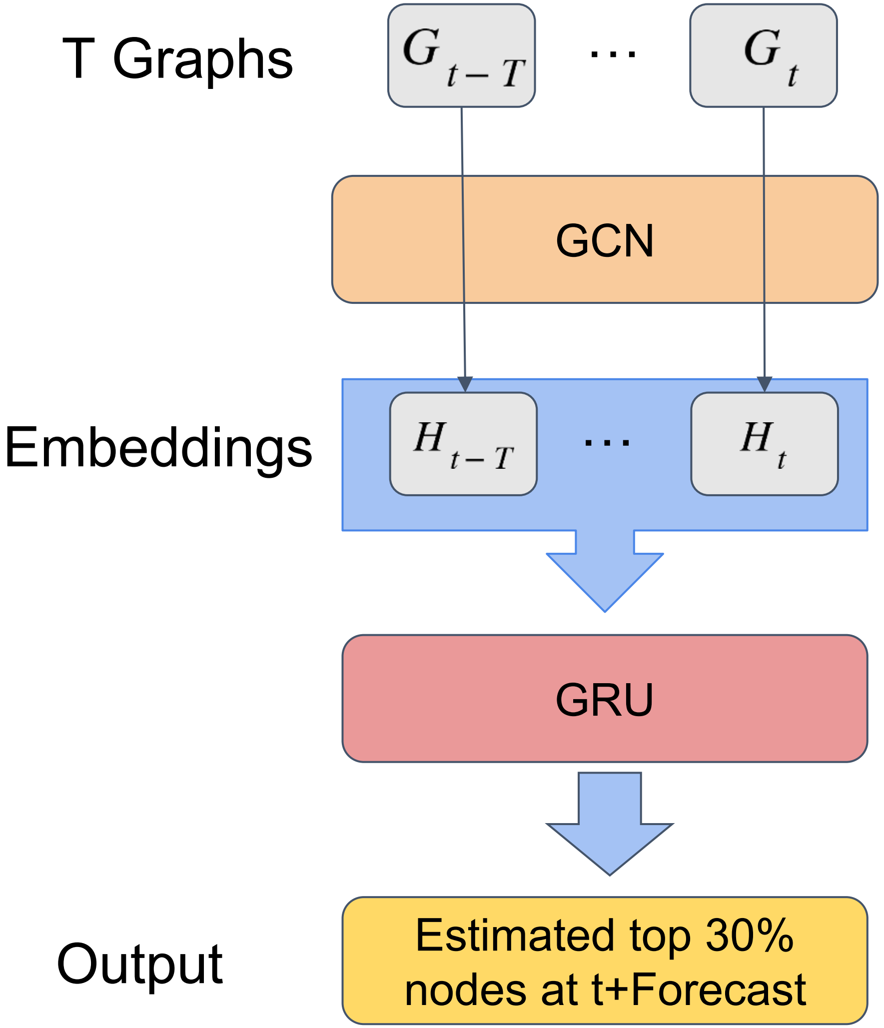
GCN on each graph (structural qnalysis)
Aggregates spatial features from neighboring nodes and encodes into embedding vectors representing each node context

GRU on each of the embedding node sequences
Processes the temporal sequence of embeddings independently for each node, capturing its evolution over time.



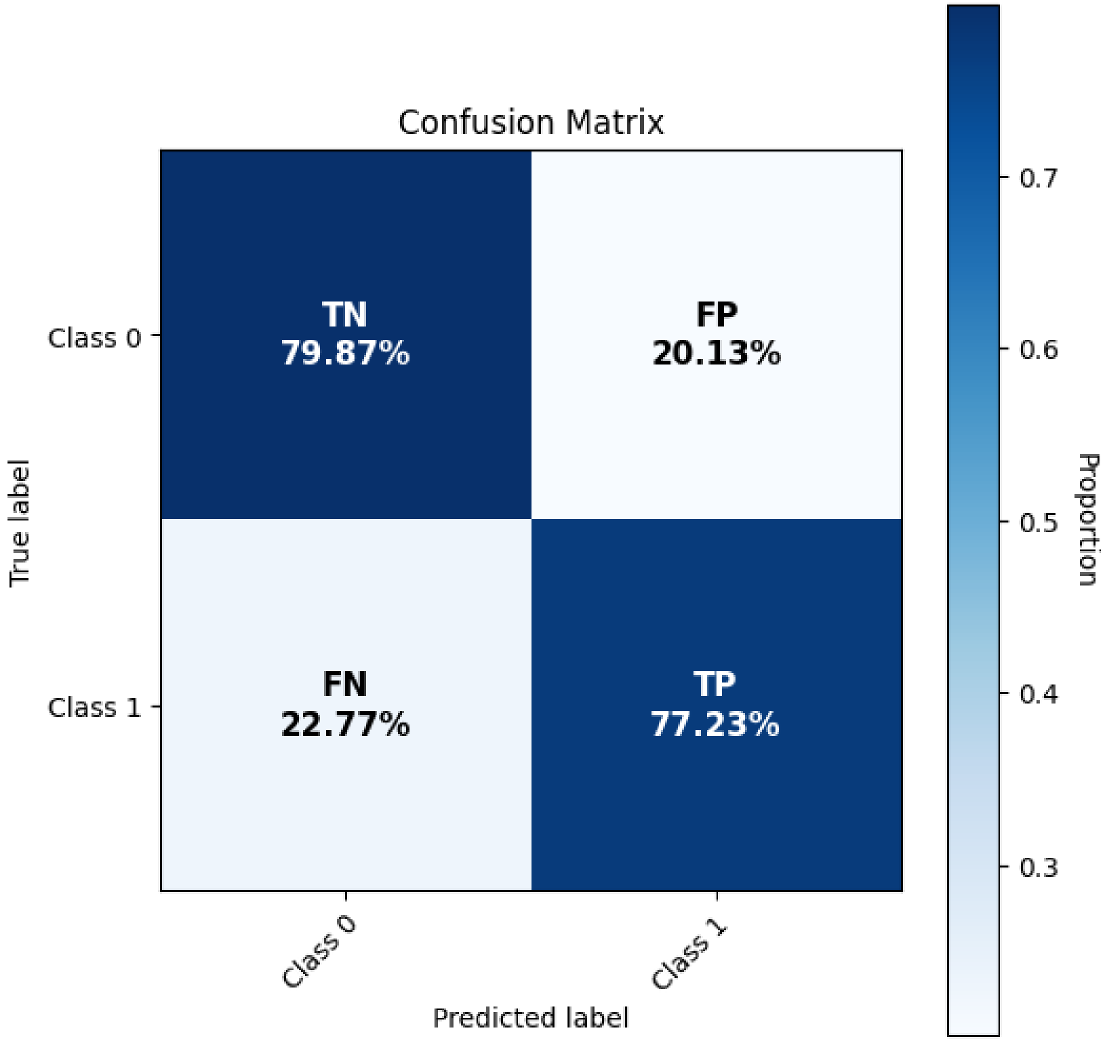
Parameter Overview

Model Parameters	
Embedding dimension	32
activation	ReLU
loss function	BCE
Training	
Sequence length (T)	4 months
Forecast	3 months
Training data	2021 - 2024
Epochs	20



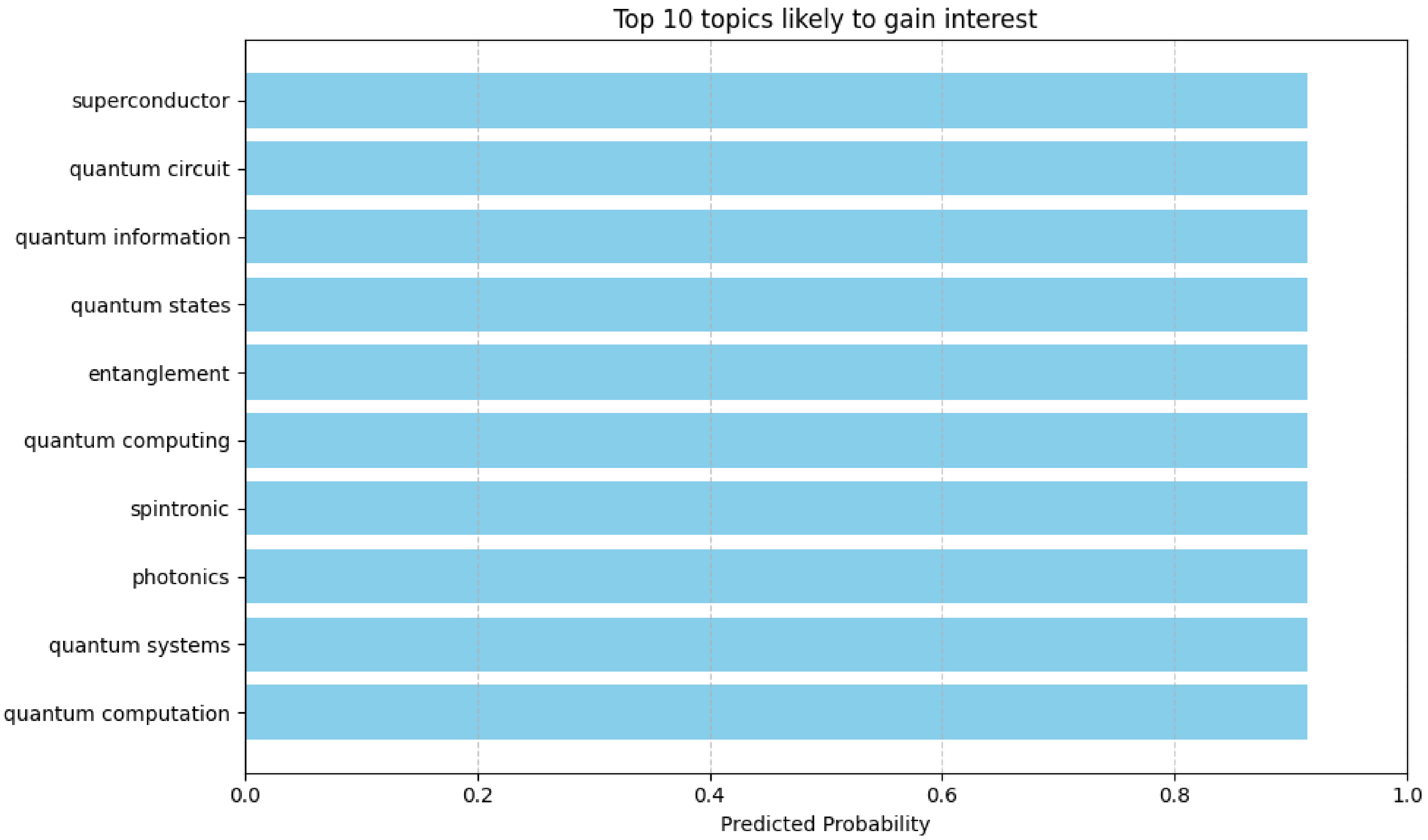
Evaluation

Naive methods	
Random predictor according to the class distribution (F1)	0.45
Majority class predictor (Weighted F1)	0.5
Our model	
Accuracy score	0.80
AUC score	0.85
F1 score	0.75



Prediction Results

Predictions using our GCN+GRU model : (3 months forecast)



Summary

✓ Keyword extraction

✓ Identification of the past trends

✓ Our GNN model seems to work for predicting emerging technology

Future Works

- Use the citation and author features
- Use other GNN architectures such as GTAN (Graph Temporal Attention Network)
- Use other model hyperparameters and emergence metrics

e.g. node emergence if score > 0.3 (arbitrary threshold)

- Apply the framework to another use case : Cyber Security in Space

Bibliography

- Cohan, A., Feldman, S., Beltagy, I., Downey,D., Weld, D.S.: Specter: Document level representation learning using citation-informed transformers. arXivpreprint arXiv:2004.07180 (2020)
- Grootendorst,M.: KeyBERT: Minimal keyword extraction with BERT. Accessed:2025-04-16(2020)
- Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
- Gu, X., Krenn, M.: Forecasting high-impact research topics via machine learning on evolving knowledge graphs. arXiv preprint arXiv:2402.08640 (2024)

Questions ?