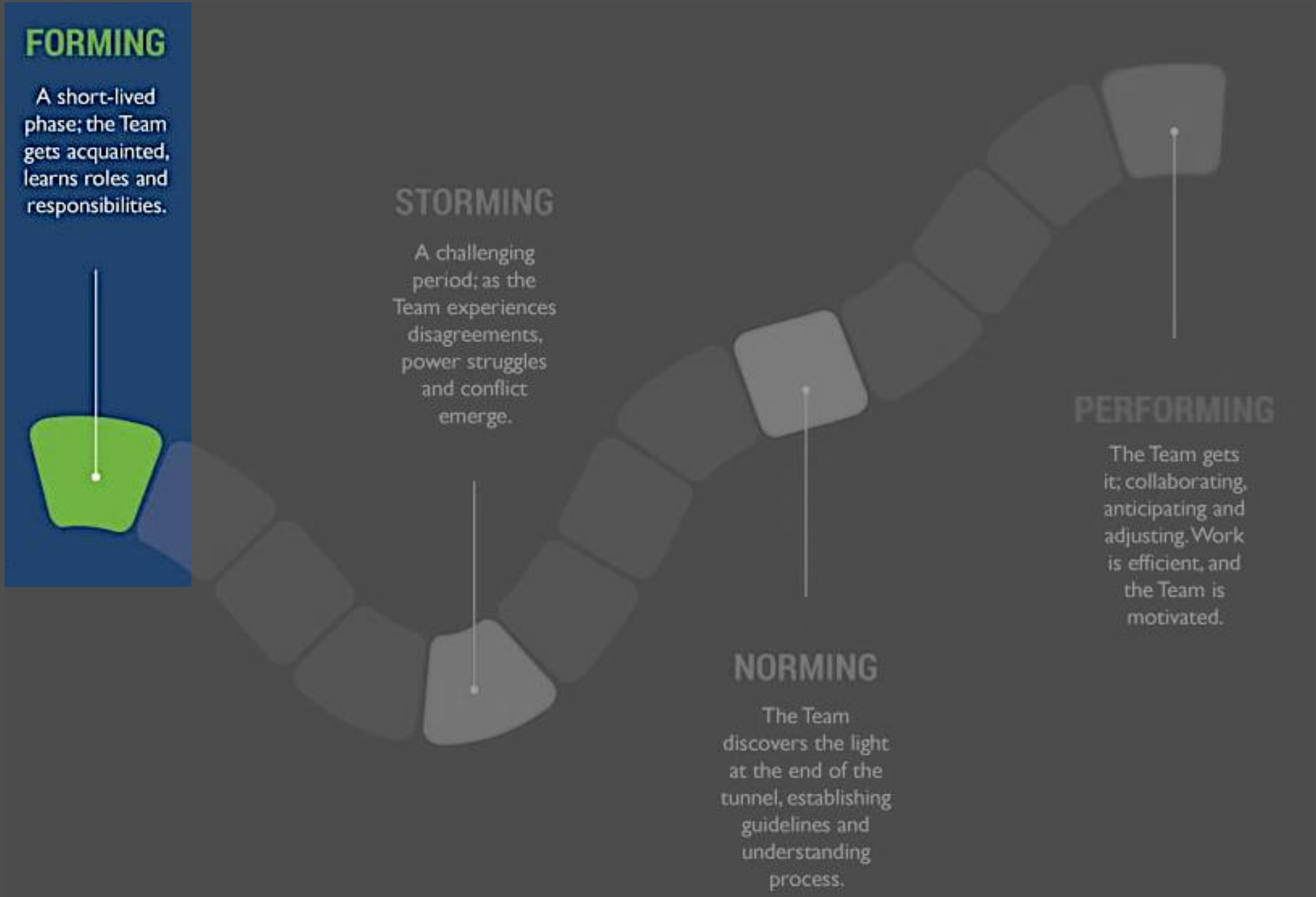
A Streaming Approach to Neural Team Formation Training

Photo: https://www.instagram.com/daviddoubilet/

5 MB hard drive being shipped by IBM - 1956.

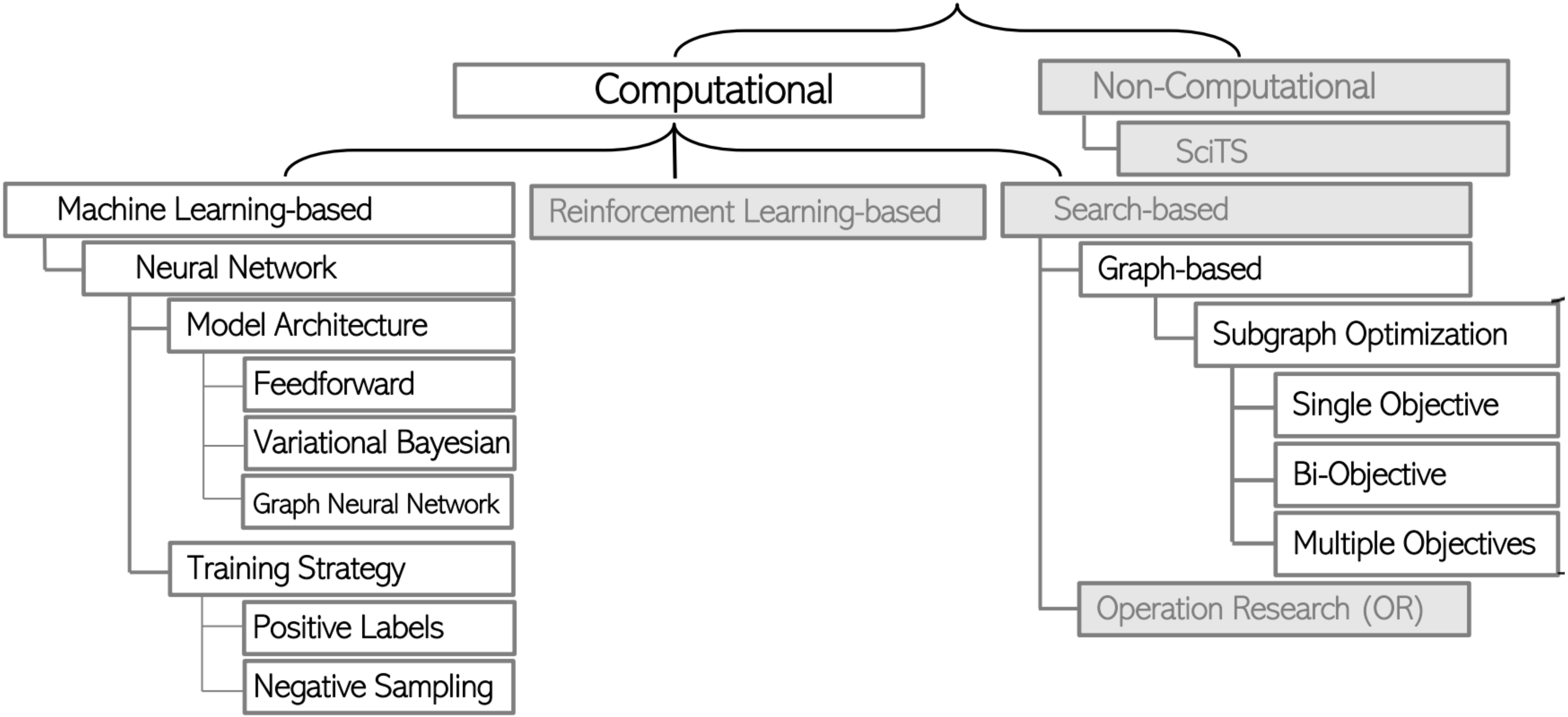


Tuckman, Bruce W. "Developmental sequence in small groups." Psychological bulletin 63.6 (1965): 384.

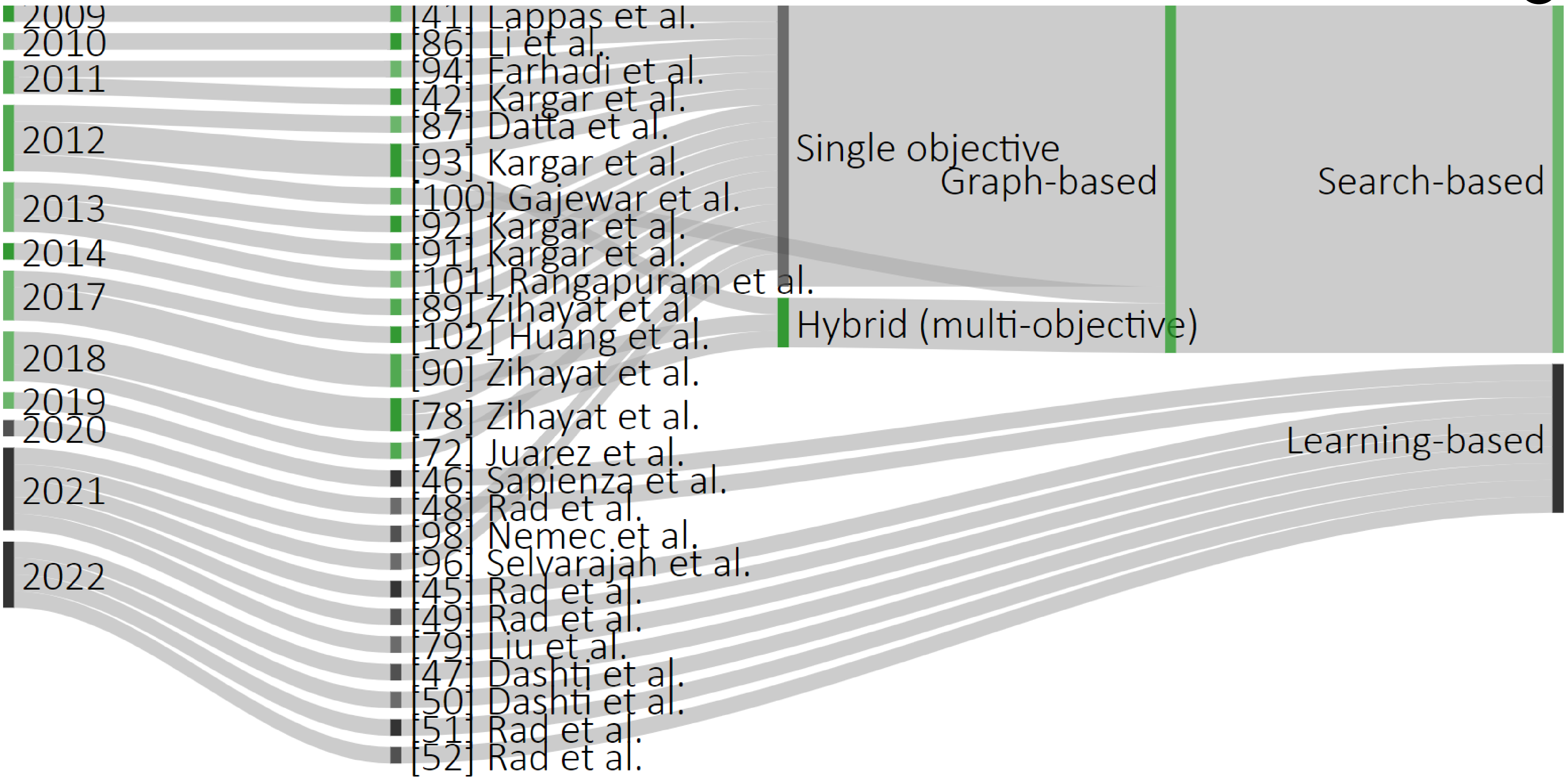
Photo: unknown

Conventionally Manual by a Human Selector:

- Large number of expert candidates
  - Different background
  - Different traits (night owls vs. early birds)
- Multitude criteria to optimize
  - Budget/Salary
  - Time/Availability
  - Communication costs
- Biases
  - Popularity
  - Gender
  - Race



Team Allocation, Team Selection, Team Composition, Team Configuration, Team Recommendation, Team Formation


2009
2010
2011
2012
2013
2014
2017
2018
2019
2020
2021
2022
[41] Lappas et al.
[86] Li et al.
[94] Farhadi et al.
[42] Kargar et al.
[87] Datta et al.
[93] Kargar et al.
[100] Gajewar et al.
[92] Kargar et al.
[91] Kargar et al.
[101] Rangapuram et al.
[89] Zihayat et al.
[102] Huang et al.
[90] Zihayat et al.
[78] Zihayat et al.
[72] Juarez et al.
[46] Sapienza et al.
[48] Rad et al.
[98] Nemec et al.
[96] Selvarajah et al.
[45] Rad et al.
[49] Rad et al.
[79] Liu et al.
[47] Dashti et al.
[50] Dashti et al.
[51] Rad et al.
[52] Rad et al.
Single objective
Hybrid (multi-objective)
Graph-based
Search-based
Learning-based

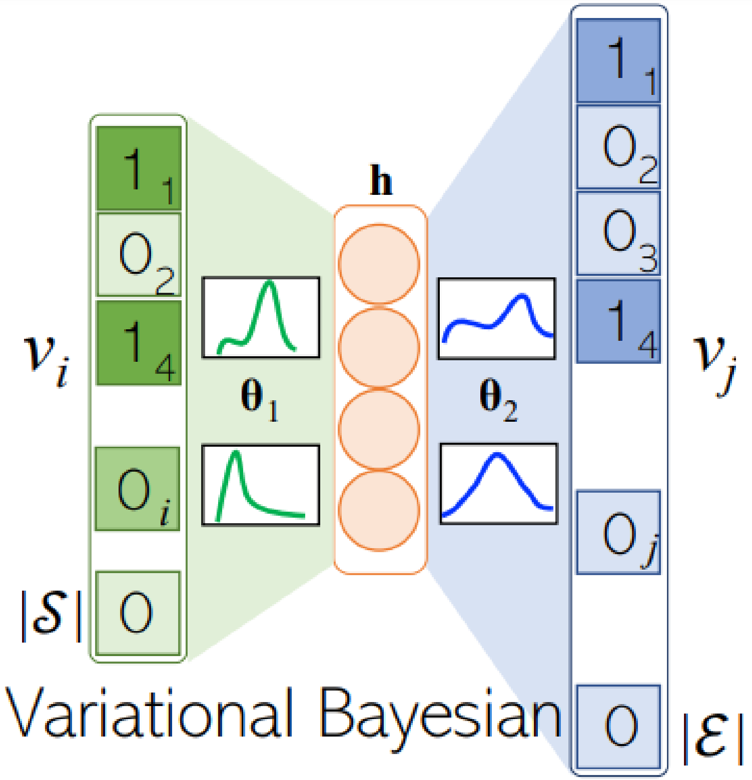
Neural Team Formation

Photo by: Julia Wimmerlin

Definition 1 (Team). *Given a set of skills $\mathcal{S} = \{i\}$ and a set of experts $\mathcal{E} = \{j\}$, a team of experts $\boldsymbol{e} \subseteq \mathcal{E}$; $\boldsymbol{e} \neq \emptyset$ that collectively cover the skill set $\boldsymbol{s} \subseteq \mathcal{S}$; $\boldsymbol{s} \neq \emptyset$ is shown by $(\boldsymbol{s}, \boldsymbol{e})$ along with its success status $y \in \{0, 1\}$. Further, $\mathcal{T} = \{(\boldsymbol{s}, \boldsymbol{e})_y : y \in \{0, 1\}\}$ indexes all previous teams.*

Definition 2 (Team Formation). *Given a subset of skills $\boldsymbol{s}$ and all teams $\mathcal{T}$, the Team Formation problem aims at identifying an optimal subset of experts $\boldsymbol{e}^*$ such that their collaboration in the predicted team $(\boldsymbol{s}, \boldsymbol{e}^*)$ is successful, that is $(\boldsymbol{s}, \boldsymbol{e}^*)_{y=1}$, while avoiding a subset of experts $\boldsymbol{e}'$ resulting in $(\boldsymbol{s}, \boldsymbol{e}')_{y=0}$. More concretely, the Team Formation problem is to find a mapping function f of parameters $\boldsymbol{\theta}$ from the powerset of skills to the powerset of experts such that $f_{\boldsymbol{\theta}} : \mathcal{P}(\mathcal{S}) \to \mathcal{P}(\mathcal{E})$, $f_{\boldsymbol{\theta}}(\boldsymbol{s}) = \boldsymbol{e}^*$.*

Definition 3 (Neural Team Formation). *Given the training set $\mathcal{T}$, Neural Team Formation estimates $f_{\boldsymbol{\theta}}(\boldsymbol{s})$ using a multi-layer neural network that learns, from $\mathcal{T}$, to map a vector representation of subset of skills $\boldsymbol{s}$, referred to as $v_{\boldsymbol{s}}$, to a vector representation of subset of experts $\boldsymbol{e}^*$, referred to as $v_{\boldsymbol{e}^*}$, by maximizing the posterior (MAP) probability of $\boldsymbol{\theta}$ in $f_{\boldsymbol{\theta}}$ over $\mathcal{T}$, that is, $\underset{\boldsymbol{\theta}}{\mathtt{argmax}}\ p(\boldsymbol{\theta}|\mathcal{T})$.*



$$\mathbf{h} = \pi(\boldsymbol{\theta}_1 v_{\mathbf{s}} + \mathbf{b}_1)$$

$$logits \rightarrow \mathbf{z} = \boldsymbol{\theta}_2 \mathbf{h} + \mathbf{b}_2$$

$$v_{\mathbf{e}^*} = \sigma(\mathbf{z})$$

$$\mathtt{argmax}_{\boldsymbol{\theta}}\ p(\boldsymbol{\theta}|\mathcal{T}) \propto p(\mathcal{T}|\boldsymbol{\theta})p(\boldsymbol{\theta}) = p(\boldsymbol{\theta}) \prod_{(\mathbf{s},\mathbf{e}^*) \in \mathcal{T}^+} p(\mathbf{e}^*|\mathbf{s}, \boldsymbol{\theta})$$

$$p(\mathbf{e}|\mathbf{s}, \boldsymbol{\theta}) = \prod_{j \in e^*} \sigma(\mathbf{z}[j]) \propto \sum_{j \in e^*} \log \sigma(\mathbf{z}[j])$$

Definition 1 (Team). *Given a set of skills $\mathcal{S} = \{i\}$ and a set of experts $\mathcal{E} = \{j\}$, a team of experts $\boldsymbol{e} \subseteq \mathcal{E}$; $\boldsymbol{e} \neq \emptyset$ that collectively cover the skill set $\boldsymbol{s} \subseteq \mathcal{S}$; $\boldsymbol{s} \neq \emptyset$ is shown by $(\boldsymbol{s}, \boldsymbol{e})$ along with its success status $y \in \{0, 1\}$. Further, $\mathcal{T} = \{(\boldsymbol{s}, \boldsymbol{e})_y : y \in \{0, 1\}\}$ indexes all previous teams.*

Definition 2 (Team Formation). *Given a subset of skills $\boldsymbol{s}$ and all teams $\mathcal{T}$, the Team Formation problem aims at identifying an optimal subset of experts $\boldsymbol{e}^*$ such that their collaboration in the predicted team $(\boldsymbol{s}, \boldsymbol{e}^*)$ is successful, that is $(\boldsymbol{s}, \boldsymbol{e}^*)_{y=1}$, while avoiding a subset of experts $\boldsymbol{e}'$ resulting in $(\boldsymbol{s}, \boldsymbol{e}')_{y=0}$. More concretely, the Team Formation problem is to find a mapping function f of parameters $\boldsymbol{\theta}$ from the powerset of skills to the powerset of experts such that $f_{\boldsymbol{\theta}} : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{P}(\mathcal{E})$, $f_{\boldsymbol{\theta}}(\boldsymbol{s}) = \boldsymbol{e}^*$.*

Definition 3 (Neural Team Formation). *Given the training set $\mathcal{T}$, Neural Team Formation estimates $f_{\boldsymbol{\theta}}(\boldsymbol{s})$ using a multi-layer neural network that learns, from $\mathcal{T}$, to map a vector representation of subset of skills $\boldsymbol{s}$, referred to as $v_{\boldsymbol{s}}$, to a vector representation of subset of experts $\boldsymbol{e}^*$, referred to as $v_{\boldsymbol{e}^*}$, by maximizing the posterior (MAP) probability of $\boldsymbol{\theta}$ in $f_{\boldsymbol{\theta}}$ over $\mathcal{T}$, that is, $\underset{\boldsymbol{\theta}}{\text{argmax}}\ p(\boldsymbol{\theta}|\mathcal{T})$.*

$\mathbf{h} = \pi(\theta_1 v_{\mathbf{s}} + \mathbf{b}_1)$

logits $\rightarrow \mathbf{z} = \theta_2 \mathbf{h} + \mathbf{b}_2$

$v_{\mathbf{e}^*} = \sigma(\mathbf{z})$

$$\text{argmax}_\theta\, p(\theta|\mathcal{T}) \propto p(\mathcal{T}|\theta)p(\theta) = p(\theta) \prod_{(\mathbf{s},\mathbf{e}^*)\in\mathcal{T}^+} p(\mathbf{e}^*|\mathbf{s}, \theta)$$

$$p(\mathbf{e}|\mathbf{s}, \theta) = \prod_{j\in e^*} \sigma(\mathbf{z}[j]) \propto \sum_{j\in e^*} \log \sigma(\mathbf{z}[j])$$

Definition 1 (Team). *Given a set of skills $\mathcal{S} = \{i\}$ and a set of experts $\mathcal{E} = \{j\}$, a team of experts $\mathbf{e} \subseteq \mathcal{E}$; $\mathbf{e} \neq \emptyset$ that collectively cover the skill set $\mathbf{s} \subseteq \mathcal{S}$; $\mathbf{s} \neq \emptyset$ is shown by $(\mathbf{s}, \mathbf{e})$ along with its success status $y \in \{0, 1\}$. Further, $\mathcal{T} = \{(\mathbf{s}, \mathbf{e})_y : y \in \{0, 1\}\}$ indexes all previous teams.*

Definition 2 (Team Formation). *Given a subset of skills $\mathbf{s}$ and all teams $\mathcal{T}$, the Team Formation problem aims at identifying an optimal subset of experts $\mathbf{e}^*$ such that their collaboration in the predicted team $(\mathbf{s}, \mathbf{e}^*)$ is successful, that is $(\mathbf{s}, \mathbf{e}^*)_{y=1}$, while avoiding a subset of experts $\mathbf{e}'$ resulting in $(\mathbf{s}, \mathbf{e}')_{y=0}$. More concretely, the Team Formation problem is to find a mapping function f of parameters $\boldsymbol{\theta}$ from the powerset of skills to the powerset of experts such that $f_{\boldsymbol{\theta}} : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{P}(\mathcal{E}), f_{\boldsymbol{\theta}}(\mathbf{s}) = \mathbf{e}^*$.*

Definition 3 (Neural Team Formation). *Given the training set $\mathcal{T}$, Neural Team Formation estimates $f_{\boldsymbol{\theta}}(\mathbf{s})$ using a multi-layer neural network that learns, from $\mathcal{T}$, to map a vector representation of subset of skills $\mathbf{s}$, referred to as $v_{\mathbf{s}}$, to a vector representation of subset of experts $\mathbf{e}^*$, referred to as $v_{\mathbf{e}^*}$, by maximizing the posterior (MAP) probability of $\boldsymbol{\theta}$ in $f_{\boldsymbol{\theta}}$ over $\mathcal{T}$, that is, $\underset{\boldsymbol{\theta}}{\texttt{argmax}}\ p(\boldsymbol{\theta}|\mathcal{T})$.*

$$\mathbf{h} = \pi(\theta_1 v_{\mathbf{s}} + \mathbf{b}_1)$$

$$logits \rightarrow \mathbf{z} = \theta_2 \mathbf{h} + \mathbf{b}_2$$

$$v_{\mathbf{e}^*} = \sigma(\mathbf{z})$$

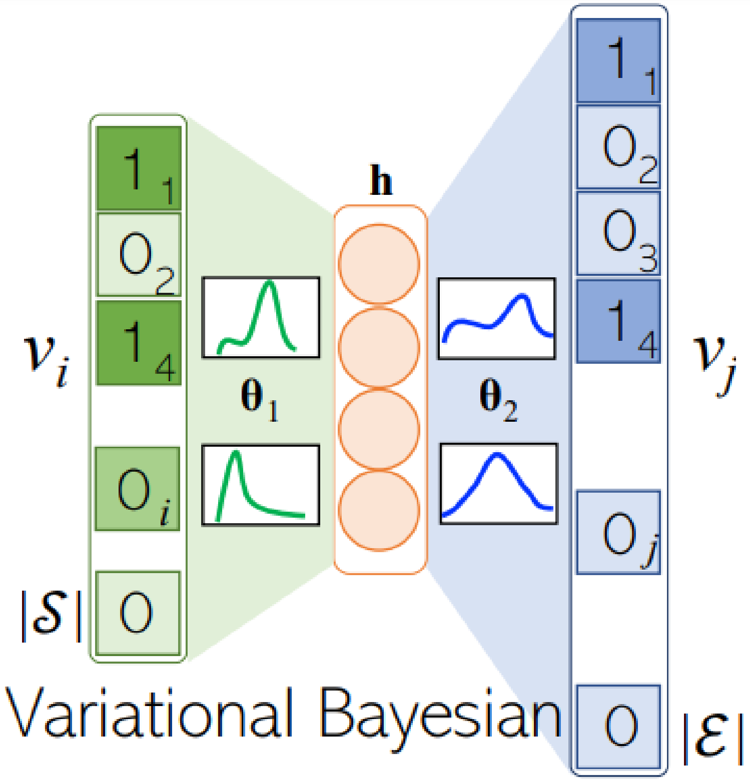
$$\mathrm{argmax}_\theta\ p(\theta|\mathcal{T}) \propto p(\mathcal{T}|\theta)p(\theta) = p(\theta) \prod_{(\mathbf{s},\mathbf{e}^*) \in \mathcal{T}^+} p(\mathbf{e}^*|\mathbf{s}, \theta)$$

$$p(\mathbf{e}|\mathbf{s}, \theta) = \prod_{j \in e^*} \sigma(\mathbf{z}[j]) \propto \sum_{j \in e^*} \log \sigma(\mathbf{z}[j])$$

Definition 1 (Team). *Given a set of skills $\mathcal{S} = \{i\}$ and a set of experts $\mathcal{E} = \{j\}$, a team of experts $\mathbf{e} \subseteq \mathcal{E}$; $\mathbf{e} \neq \emptyset$ that collectively cover the skill set $\mathbf{s} \subseteq \mathcal{S}$; $\mathbf{s} \neq \emptyset$ is shown by $(\mathbf{s}, \mathbf{e})$ along with its success status $y \in \{0, 1\}$. Further, $\mathcal{T} = \{(\mathbf{s}, \mathbf{e})_y : y \in \{0, 1\}\}$ indexes all previous teams.*

Definition 2 (Team Formation). *Given a subset of skills $\mathbf{s}$ and all teams $\mathcal{T}$, the Team Formation problem aims at identifying an optimal subset of experts $\mathbf{e}^*$ such that their collaboration in the predicted team $(\mathbf{s}, \mathbf{e}^*)$ is successful, that is $(\mathbf{s}, \mathbf{e}^*)_{y=1}$, while avoiding a subset of experts $\mathbf{e}'$ resulting in $(\mathbf{s}, \mathbf{e}')_{y=0}$. More concretely, the Team Formation problem is to find a mapping function f of parameters $\boldsymbol{\theta}$ from the powerset of skills to the powerset of experts such that $f_{\boldsymbol{\theta}} : \mathcal{P}(\mathcal{S}) \to \mathcal{P}(\mathcal{E})$, $f_{\boldsymbol{\theta}}(\mathbf{s}) = \mathbf{e}^*$.*

Definition 3 (Neural Team Formation). *Given the training set $\mathcal{T}$, Neural Team Formation estimates $f_{\boldsymbol{\theta}}(\mathbf{s})$ using a multi-layer neural network that learns, from $\mathcal{T}$, to map a vector representation of subset of skills $\mathbf{s}$, referred to as $v_{\mathbf{s}}$, to a vector representation of subset of experts $\mathbf{e}^*$, referred to as $v_{\mathbf{e}^*}$, by maximizing the posterior (MAP) probability of $\boldsymbol{\theta}$ in $f_{\boldsymbol{\theta}}$ over $\mathcal{T}$, that is, $\underset{\boldsymbol{\theta}}{\texttt{argmax}}\ p(\boldsymbol{\theta}|\mathcal{T})$.*



$$\mathbf{h} = \pi(\boldsymbol{\theta}_1 v_{\mathbf{s}} + \mathbf{b}_1)$$

$$logits \to \mathbf{z} = \boldsymbol{\theta}_2 \mathbf{h} + \mathbf{b}_2$$

$$v_{\mathbf{e}^*} = \sigma(\mathbf{z})$$

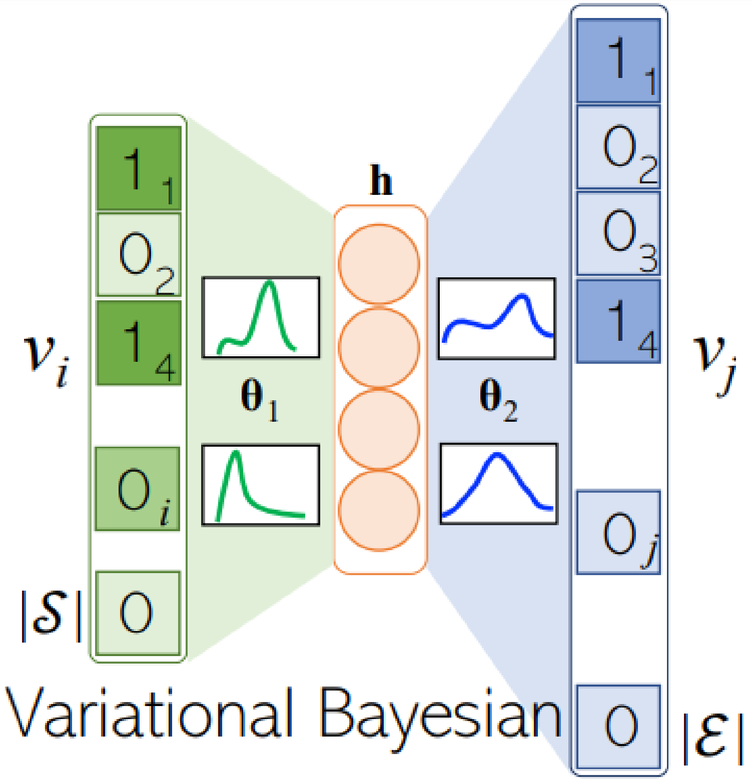
$$\texttt{argmax}_{\theta}\ p(\theta|\mathcal{T}) \propto p(\mathcal{T}|\theta)p(\theta) = p(\theta) \prod_{(\mathbf{s},\mathbf{e}^*) \in \mathcal{T}^+} p(\mathbf{e}^*|\mathbf{s}, \theta)$$

$$p(\mathbf{e}|\mathbf{s}, \theta) = \prod_{j \in e^*} \sigma(\mathbf{z}[j]) \propto \sum_{j \in e^*} \log \sigma(\mathbf{z}[j])$$

Definition 1 (Team). *Given a set of skills* $\mathcal{S} = \{i\}$ *and a set of experts* $\mathcal{E} = \{j\}$*, a team of experts* $\mathbf{e} \subseteq \mathcal{E}$*;* $\mathbf{e} \neq \emptyset$ *that collectively cover the skill set* $\mathbf{s} \subseteq \mathcal{S}$*;* $\mathbf{s} \neq \emptyset$ *is shown by* $(\mathbf{s}, \mathbf{e})$ *along with its success status* $y \in \{0, 1\}$*. Further,* $\mathcal{T} = \{(\mathbf{s}, \mathbf{e})_y : y \in \{0, 1\}\}$ *indexes all previous teams.*

Definition 2 (Team Formation). *Given a subset of skills* $\mathbf{s}$ *and all teams* $\mathcal{T}$*, the Team Formation problem aims at identifying an optimal subset of experts* $\mathbf{e}^*$ *such that their collaboration in the predicted team* $(\mathbf{s}, \mathbf{e}^*)$ *is successful, that is* $(\mathbf{s}, \mathbf{e}^*)_{y=1}$*, while avoiding a subset of experts* $\mathbf{e}'$ *resulting in* $(\mathbf{s}, \mathbf{e}')_{y=0}$*. More concretely, the Team Formation problem is to find a mapping function* f *of parameters* $\boldsymbol{\theta}$ *from the powerset of skills to the powerset of experts such that* $f_{\boldsymbol{\theta}} : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{P}(\mathcal{E})$*,* $f_{\boldsymbol{\theta}}(\mathbf{s}) = \mathbf{e}^*$*.*

Definition 3 (Neural Team Formation). *Given the training set* $\mathcal{T}$*, Neural Team Formation estimates* $f_{\boldsymbol{\theta}}(\mathbf{s})$ *using a multi-layer neural network that learns, from* $\mathcal{T}$*, to map a vector representation of subset of skills* $\mathbf{s}$*, referred to as* $v_{\mathbf{s}}$*, to a vector representation of subset of experts* $\mathbf{e}^*$*, referred to as* $v_{\mathbf{e}^*}$*, by maximizing the posterior (MAP) probability of* $\boldsymbol{\theta}$ *in* $f_{\boldsymbol{\theta}}$ *over* $\mathcal{T}$*, that is,* $\underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ p(\boldsymbol{\theta}|\mathcal{T})$*.*



$$\mathbf{h} = \pi(\boldsymbol{\theta}_1 v_{\mathbf{s}} + \mathbf{b}_1)$$

$$logits \rightarrow \mathbf{z} = \boldsymbol{\theta}_2 \mathbf{h} + \mathbf{b}_2$$

$$v_{\mathbf{e}^*} = \sigma(\mathbf{z})$$

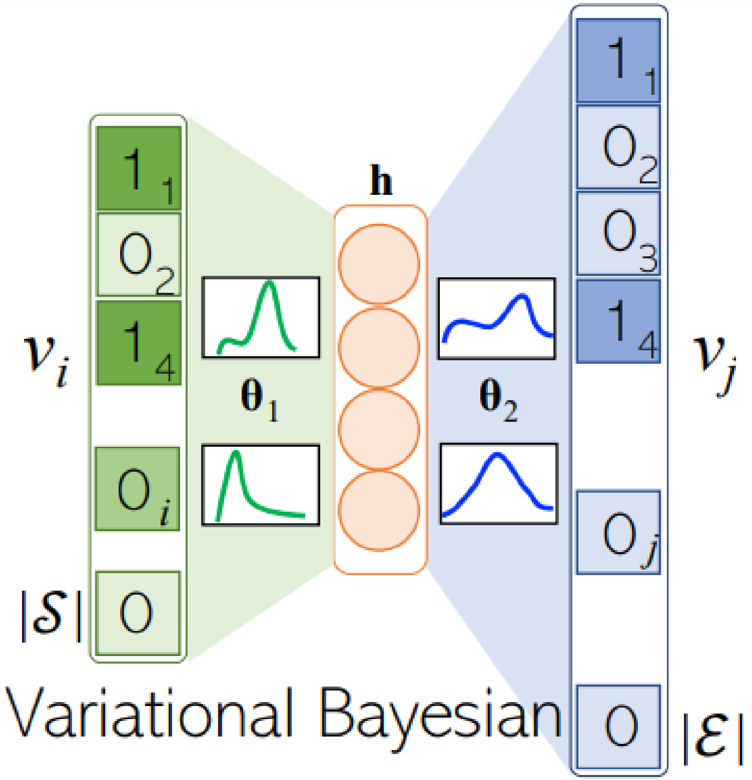
$$\operatorname{argmax}_{\boldsymbol{\theta}}\ p(\boldsymbol{\theta}|\mathcal{T}) \propto p(\mathcal{T}|\boldsymbol{\theta})p(\boldsymbol{\theta}) = p(\boldsymbol{\theta}) \prod_{(\mathbf{s},\mathbf{e}^*) \in \mathcal{T}^+} p(\mathbf{e}^*|\mathbf{s}, \boldsymbol{\theta})$$

$$p(\mathbf{e}|\mathbf{s}, \theta) = \prod_{j \in e^*} \sigma(\mathbf{z}[j]) \propto \sum_{j \in e^*} \log \sigma(\mathbf{z}[j])$$

Definition 1 (Team). *Given a set of skills $\mathcal{S} = \{i\}$ and a set of experts $\mathcal{E} = \{j\}$, a team of experts $\mathbf{e} \subseteq \mathcal{E}$; $\mathbf{e} \neq \emptyset$ that collectively cover the skill set $\mathbf{s} \subseteq \mathcal{S}$; $\mathbf{s} \neq \emptyset$ is shown by $(\mathbf{s}, \mathbf{e})$ along with its success status $y \in \{0, 1\}$. Further, $\mathcal{T} = \{(\mathbf{s}, \mathbf{e})_y : y \in \{0, 1\}\}$ indexes all previous teams.*

Definition 2 (Team Formation). *Given a subset of skills $\mathbf{s}$ and all teams $\mathcal{T}$, the Team Formation problem aims at identifying an optimal subset of experts $\mathbf{e}^*$ such that their collaboration in the predicted team $(\mathbf{s}, \mathbf{e}^*)$ is successful, that is $(\mathbf{s}, \mathbf{e}^*)_{y=1}$, while avoiding a subset of experts $\mathbf{e}'$ resulting in $(\mathbf{s}, \mathbf{e}')_{y=0}$. More concretely, the Team Formation problem is to find a mapping function f of parameters $\boldsymbol{\theta}$ from the powerset of skills to the powerset of experts such that $f_{\boldsymbol{\theta}} : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{P}(\mathcal{E})$, $f_{\boldsymbol{\theta}}(\mathbf{s}) = \mathbf{e}^*$.*

Definition 3 (Neural Team Formation). *Given the training set $\mathcal{T}$, Neural Team Formation estimates $f_{\boldsymbol{\theta}}(\mathbf{s})$ using a multi-layer neural network that learns, from $\mathcal{T}$, to map a vector representation of subset of skills $\mathbf{s}$, referred to as $v_{\mathbf{s}}$, to a vector representation of subset of experts $\mathbf{e}^*$, referred to as $v_{\mathbf{e}^*}$, by maximizing the posterior (MAP) probability of $\boldsymbol{\theta}$ in $f_{\boldsymbol{\theta}}$ over $\mathcal{T}$, that is, $\underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ p(\boldsymbol{\theta}|\mathcal{T})$.*



$$\mathbf{h} = \pi(\boldsymbol{\theta}_1 v_{\mathbf{s}} + \mathbf{b}_1)$$

$$logits \rightarrow \mathbf{z} = \boldsymbol{\theta}_2 \mathbf{h} + \mathbf{b}_2$$

$$v_{\mathbf{e}^*} = \sigma(\mathbf{z})$$

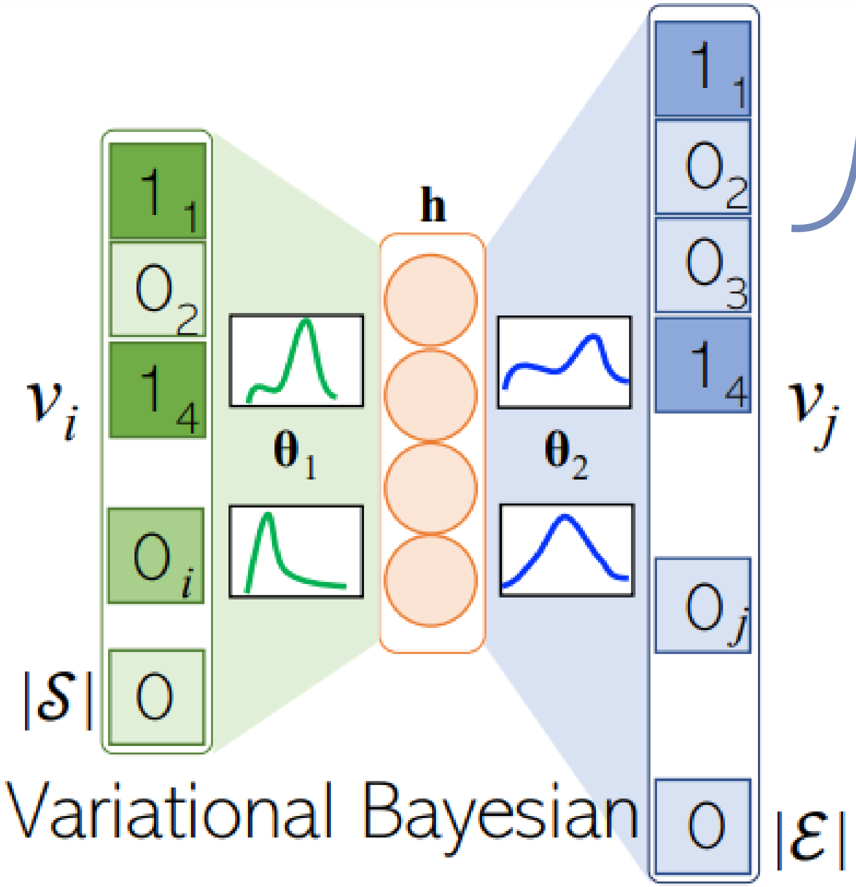
$$\operatorname{argmax}_{\boldsymbol{\theta}}\ p(\boldsymbol{\theta}|\mathcal{T}) \propto p(\mathcal{T}|\boldsymbol{\theta})p(\boldsymbol{\theta}) = p(\boldsymbol{\theta}) \prod_{(\mathbf{s},\mathbf{e}^*) \in \mathcal{T}^+} p(\mathbf{e}^*|\mathbf{s}, \boldsymbol{\theta})$$

$$p(\mathbf{e}|\mathbf{s}, \boldsymbol{\theta}) = \prod_{j \in e^*} \sigma(\mathbf{z}[j]) \propto \sum_{j \in e^*} \log \sigma(\mathbf{z}[j])$$

A Streaming Approach to Neural Team Formation Training

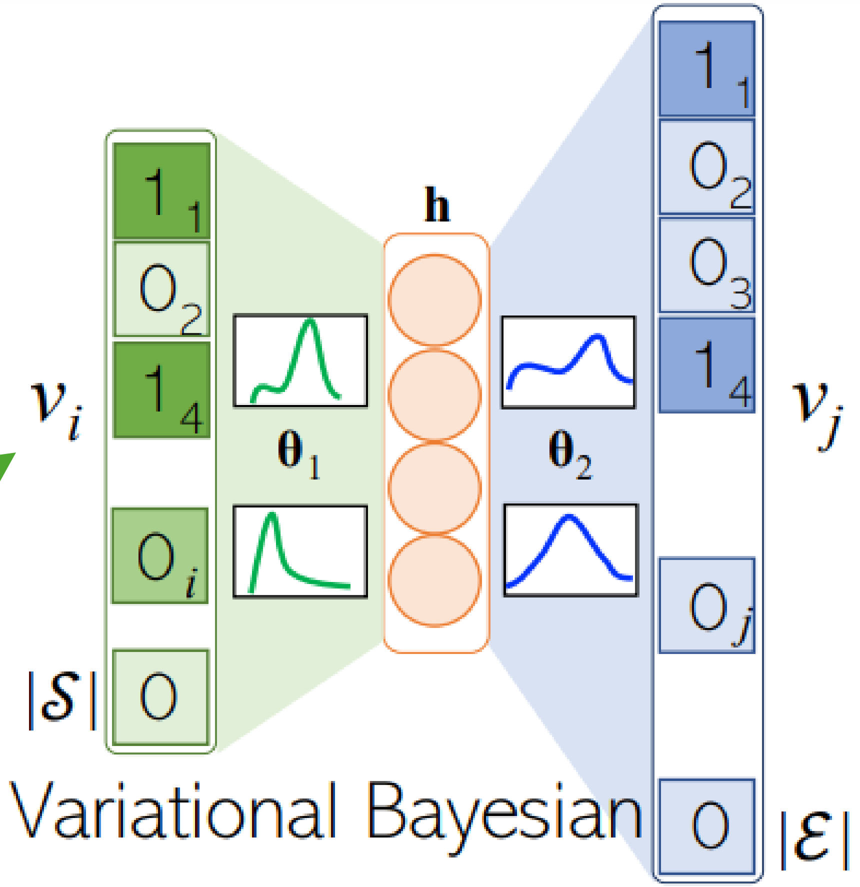
Hossein Fani[0000–0002–6033–6564], Reza Barzegar[0009–0002–2831–4143], Arman Dashti[0000–0001–9022–5403], and Mahdis Saeedi[0000–0002–6297–3794]

University of Windsor, Windsor, ON., Canada
{hfani, barzegar, vaghehd, msaeedi}@uwindsor.ca



Abstract. Predicting *future* successful teams of experts who can effectively collaborate is challenging due to the experts' temporality of skill sets, levels of expertise, and collaboration ties, which is overlooked by prior work. Specifically, state-of-the-art neural-based methods learn vector representations of experts and skills in a *static* latent space, falling short of incorporating the possible drift and variability of experts' skills and collaboration ties in time. In this paper, we propose (1) a streaming-based training strategy for neural models to capture the evolution of experts' skills and collaboration ties over time and (2) to consume time information as an additional signal to the model for predicting future successful teams. We empirically benchmark our proposed method against state-of-the-art neural team formation methods and a strong temporal recommender system on datasets from varying domains with distinct distributions of skills and experts in teams. The results demonstrate neural models that utilize our proposed training strategy excel at efficacy in terms of classification and information retrieval metrics. The codebase is available at https://github.com/fani-lab/OpeNTF/tree/ecir24.

Keywords: Neural Team Formation · Training Strategy · OpeNTF.



US1781541A

United States

Download PDF Find Prior Art Similar

Inventor: Einstein Albert, Szilard Leo

Current Assignee : Electrolux Servel Corp

Worldwide applications

1927 • US

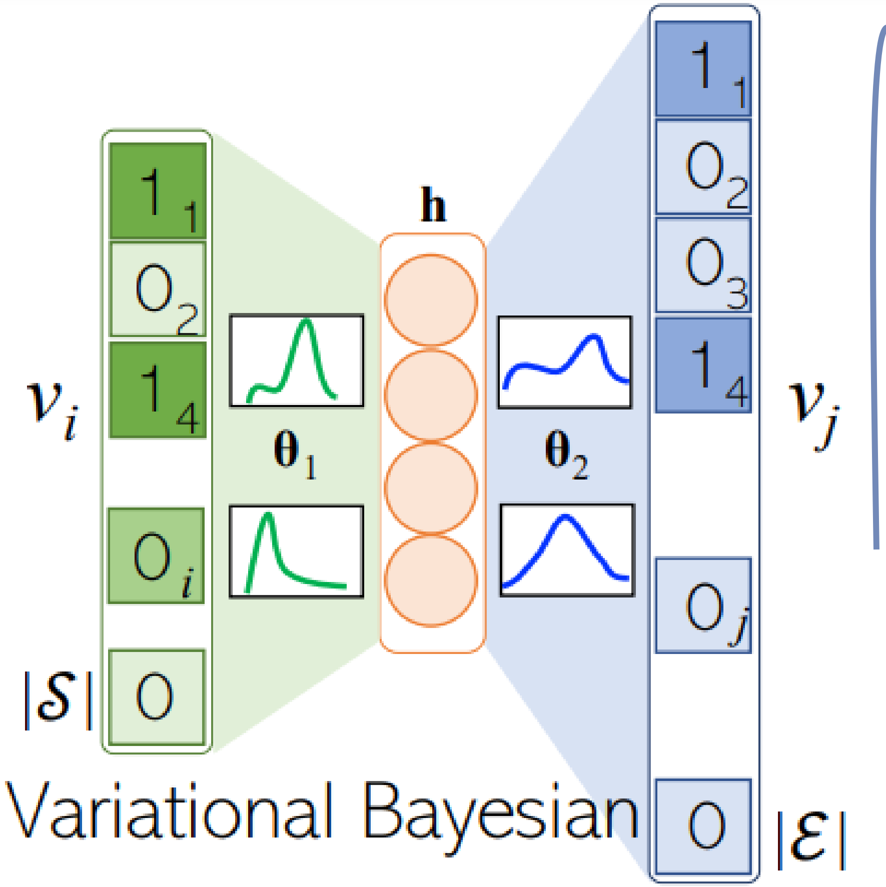
Application US240566A events

- **1927-12-16** • Application filed by Electrolux Servel Corp
- **1930-11-11** • Application granted
- **1930-11-11** • Publication of US1781541A
- **1947-11-11** • Anticipated expiration
- **Status** • Expired - Lifetime

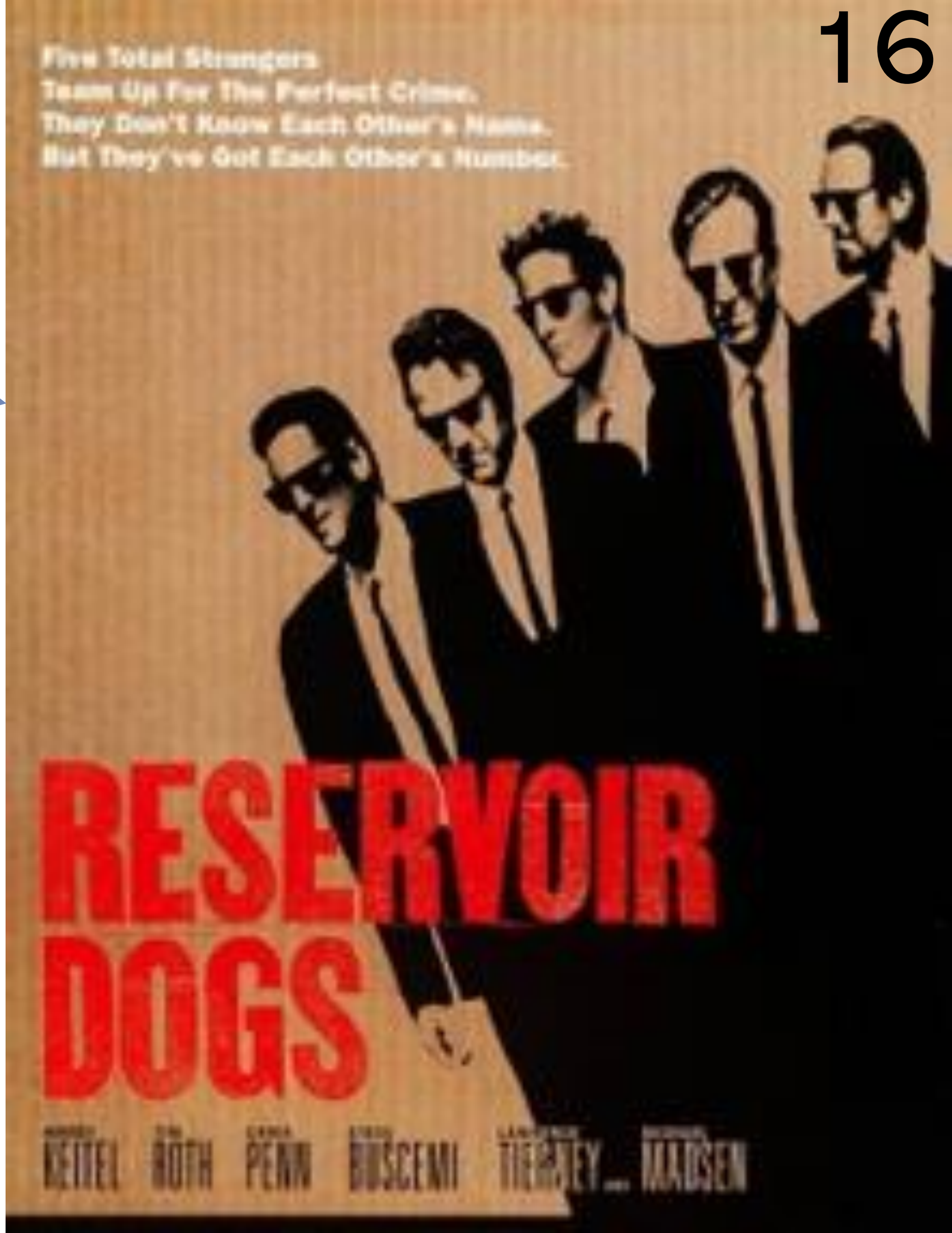
Classifications

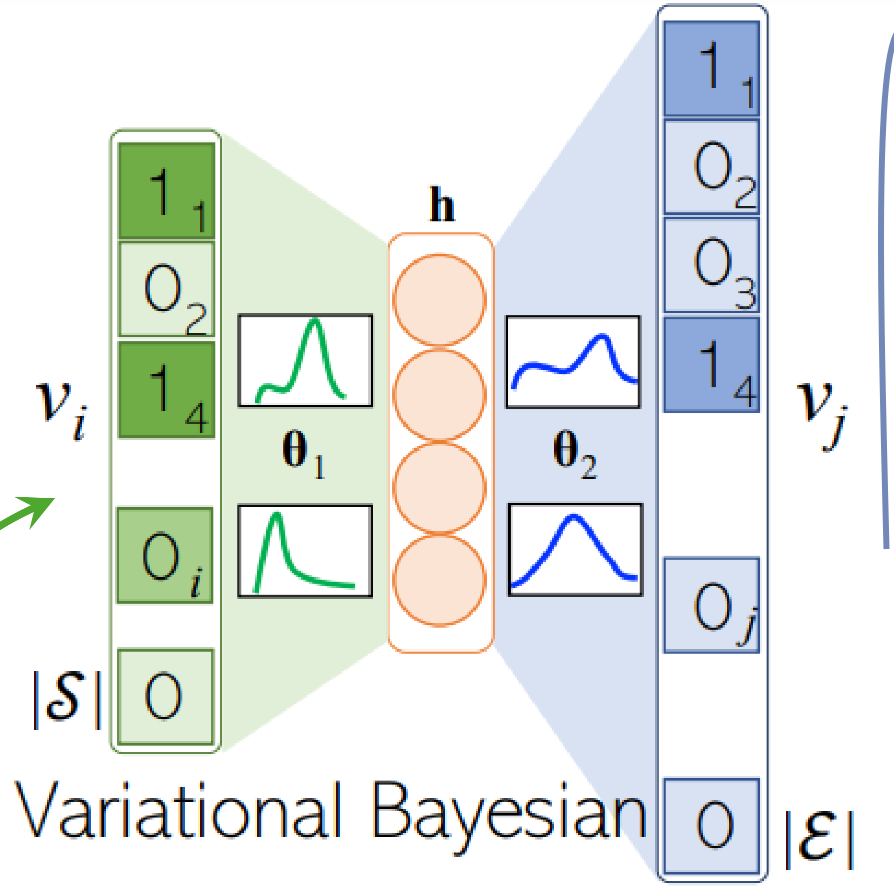
F25B15/10 Sorption machines, plants or systems, operating continuously, e.g. absorption type with inert gas

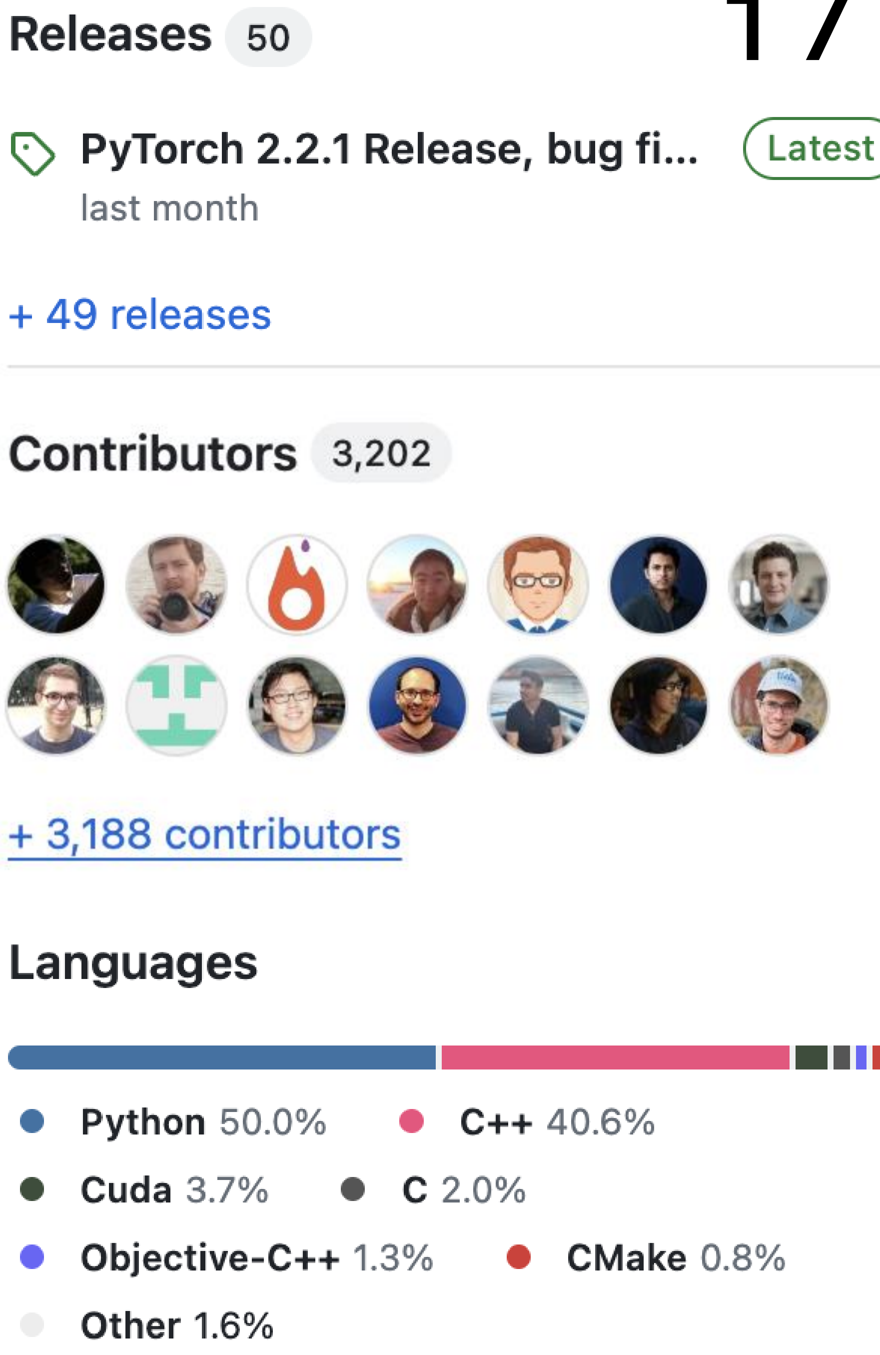
View 4 more classifications



Crime, Thriller, Neo-Noir




v_i
h
θ_1
θ_2
v_j
$|\mathcal{S}|$
$|\mathcal{E}|$
Variational Bayesian


Releases 50
PyTorch 2.2.1 Release, bug fi...
Latest
last month
+ 49 releases
Contributors 3,202
+ 3,188 contributors
Languages
Python 50.0%
C++ 40.6%
Cuda 3.7%
C 2.0%
Objective-C++ 1.3%
CMake 0.8%
Other 1.6%

Small vs. large set
- Future RQ

Small vs. large set
- Dense Representation Learning
  - GNN-based (Rad et al. SIGIR 2021)

Definition 1 (Team). *Given a set of skills $\mathcal{S} = \{i\}$ and a set of experts $\mathcal{E} = \{j\}$, a team of experts $\mathbf{e} \subseteq \mathcal{E}$; $\mathbf{e} \neq \emptyset$ that collectively cover the skill set $\mathbf{s} \subseteq \mathcal{S}$; $\mathbf{s} \neq \emptyset$ is shown by $(\mathbf{s}, \mathbf{e})$ along with its success status $y \in \{0,1\}$. Further, $\mathcal{T} = \{(\mathbf{s}, \mathbf{e})_y : y \in \{0,1\}\}$ indexes all previous teams.*

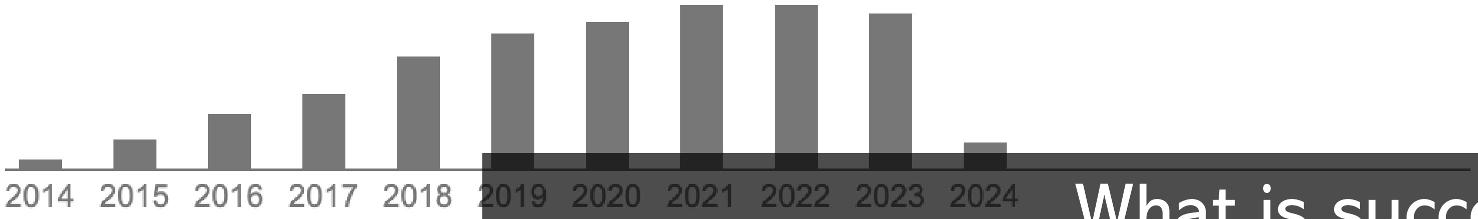
What does it mean for a team to be successful?

Challenges …

Tomas Mikolov

Efficient estimation of word representations in vector space

Authors	Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean
Publication date	2013/1/16
Journal	arXiv preprint arXiv:1301.3781
Description	We propose two novel model architectures for computing continuous vector representations of words from very large data sets. The quality of these representations is measured in a word similarity task, and the results are compared to the previously best performing techniques based on different types of neural networks. We observe large improvements in accuracy at much lower computational cost, i.e. it takes less than a day to learn high quality word vectors from a 1.6 billion words data set. Furthermore, we show that these vectors provide state-of-the-art performance on our test set for measuring syntactic and semantic word similarities.
Total citations	Cited by 40332



What is success?

Tomas Mikolov

December 13, 2023 ·

https://openreview.net/forum?id=idpCdOWtqXd60

Yesterday we received a Test of Time Award at NeurIPS for the word2vec paper from ten years ago. I'm really happy about it! I think it's the first "best paper" type of award I ever received. In fact, the original word2vec paper was rejected at the first ICLR conference in 2013 (despite the acceptance rate of around 70%), so it made me think how difficult it is for reviewers to predict future impact of research papers.

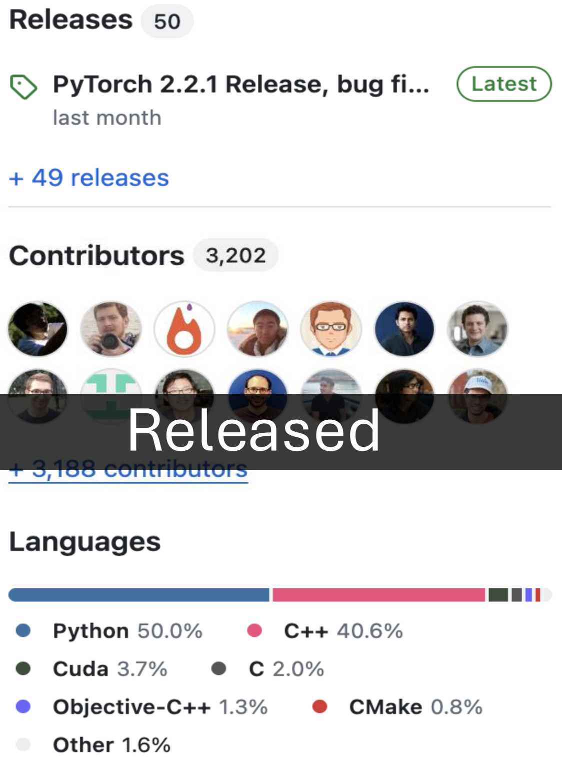
https://www.facebook.com/share/p/kXYaYaRvRCr5K2Ze

What is success?

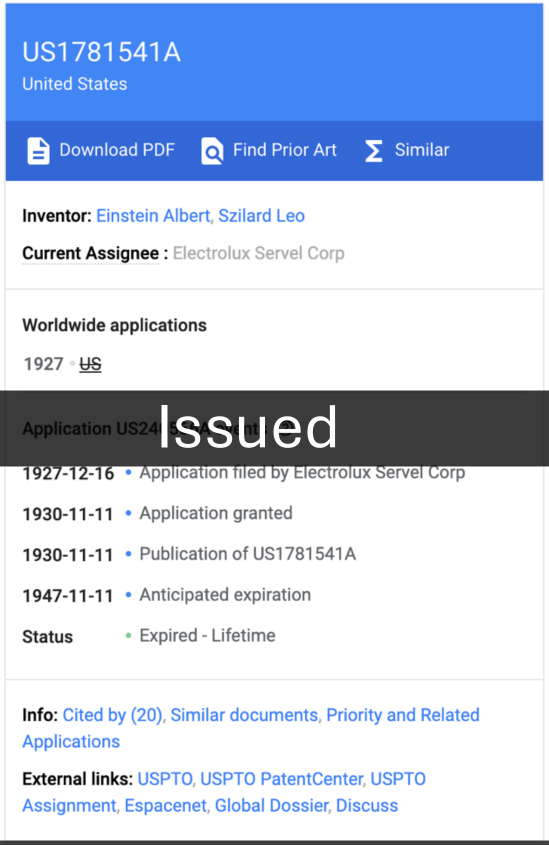
US$1.446 billion vs. no Oscar!

The Big Lebowski, 1998
Joel & Ethan Coen
Jeff Bridges, John Goodman, Steve Buscemi

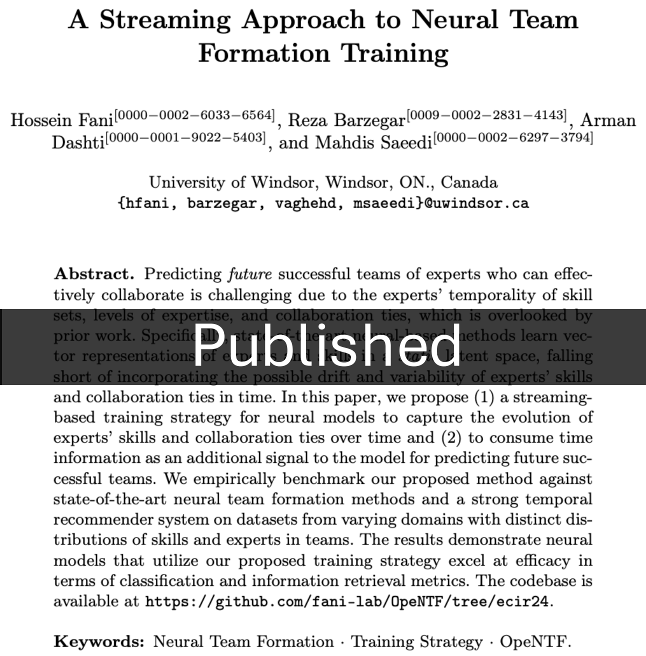
Produced



Released



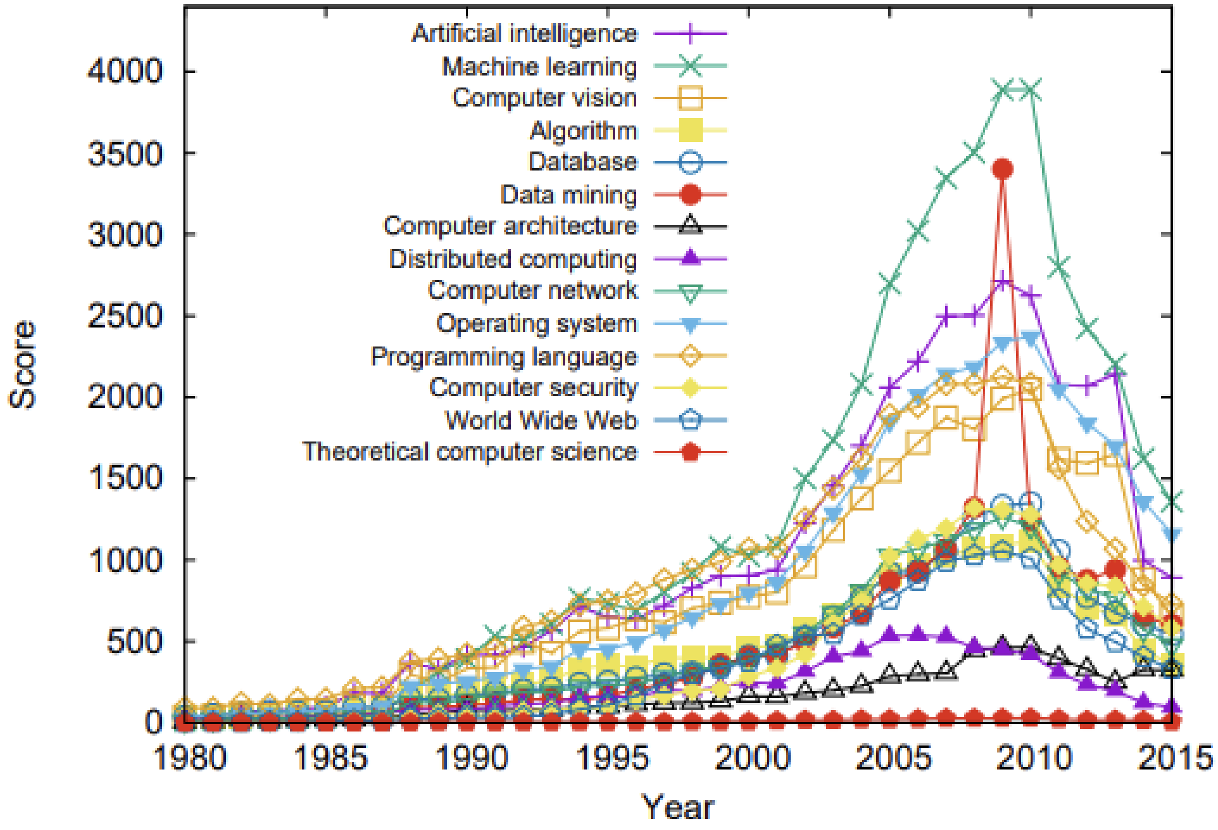
Issued



Published

Success

(a) General trend (absolute).



Effendy et al. Analysing trends in computer science research: A preliminary study using the microsoft academic graph. WWW 2017.

Operating Systems

Programming Languages

Distributed Systems

Barbara Liskov

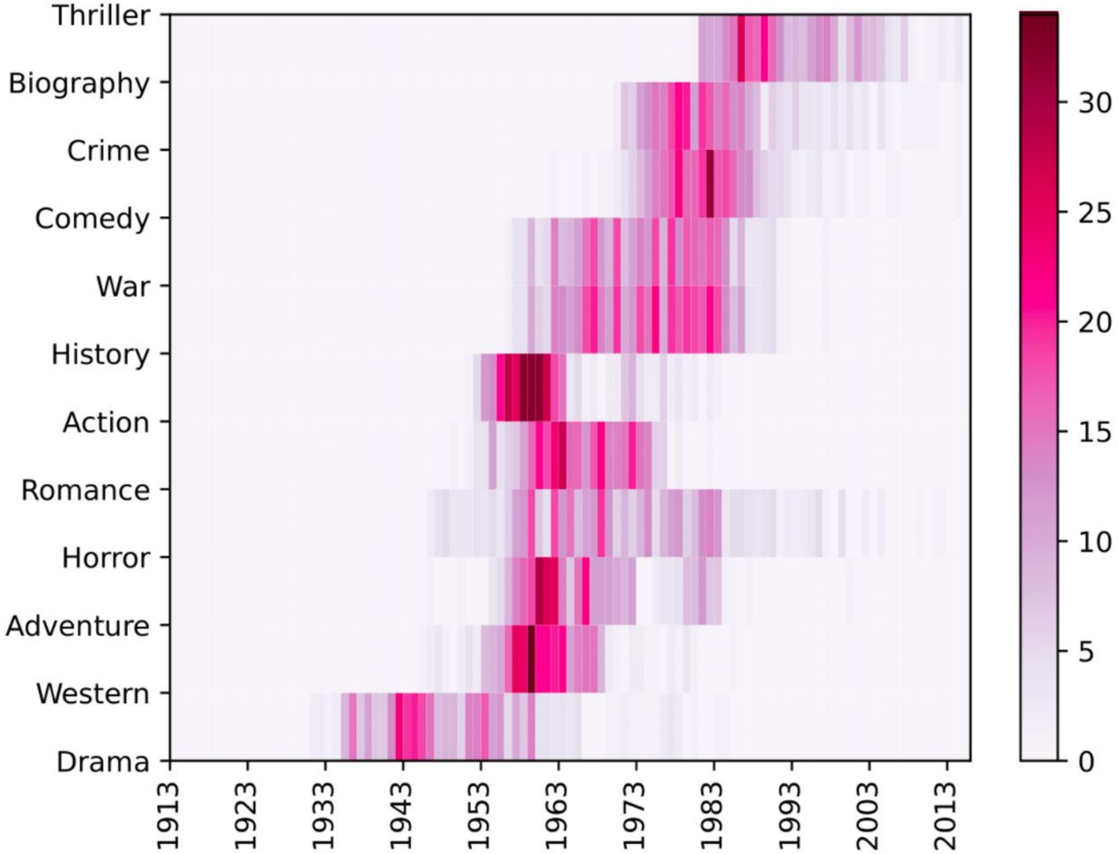
Liskov in 2010

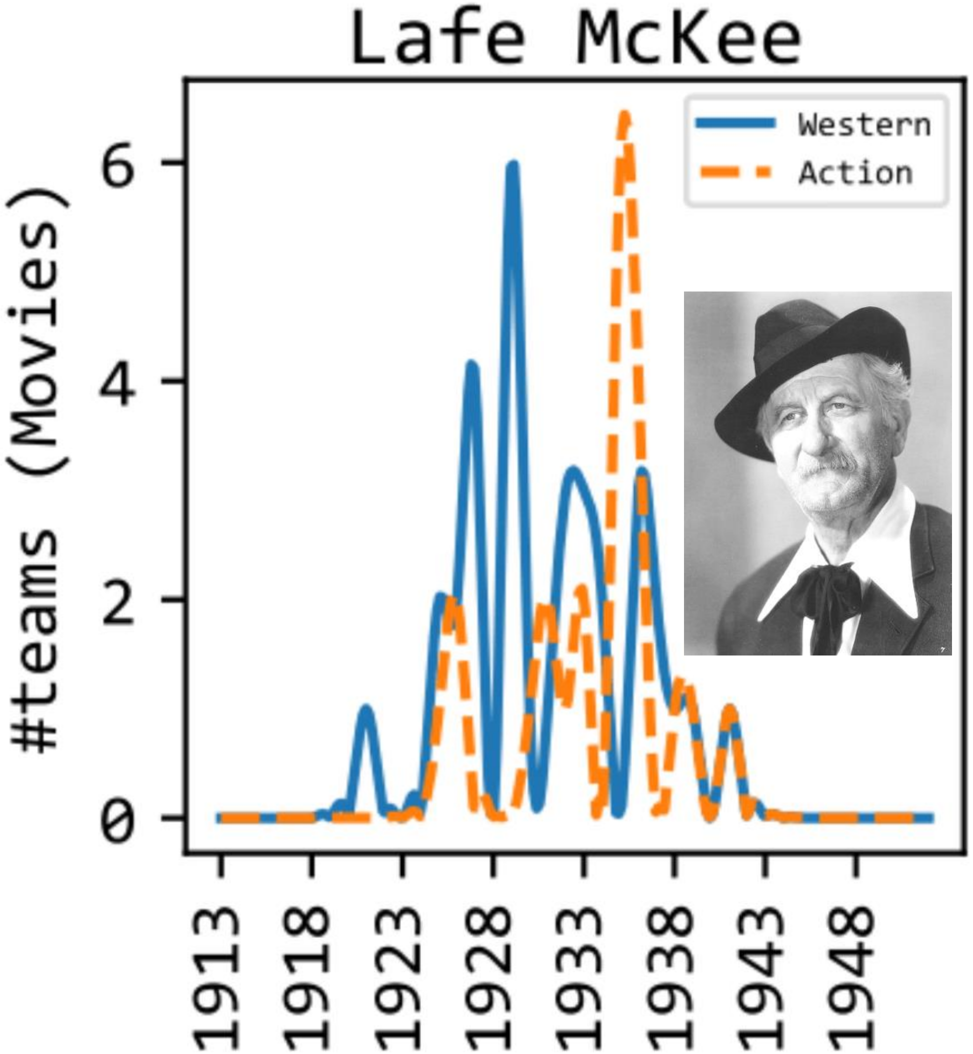
Born	Barbara Jane Huberman November 7, 1939 (age 84) Los Angeles, California, US
Alma mater	University of California, Berkeley (BA) Stanford University (PhD)
Known for	Venus (operating system) CLU Argus Thor (object-oriented database) Liskov substitution principle
Spouse	Nathan Liskov (1970–)
Children	1
	A. M. Turing Award (2008) Computer Pioneer Award (2018)

Scientific career

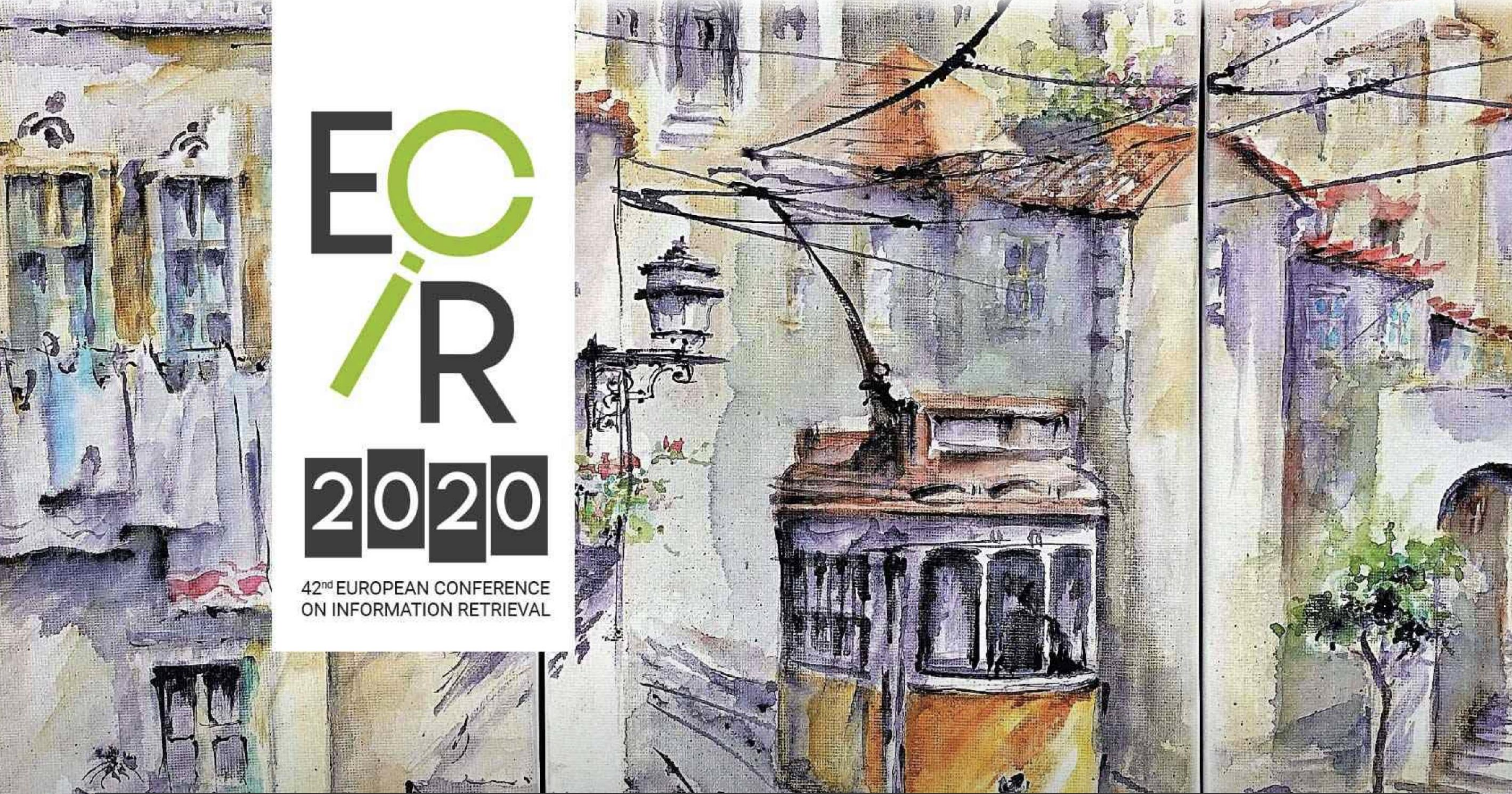
Fields	Computer science

Temporal Evolutions in Skills & Expert's Skills





Temporal Evolutions in Skills & Expert’s Skills



ECIR 2020

42nd EUROPEAN CONFERENCE ON INFORMATION RETRIEVAL

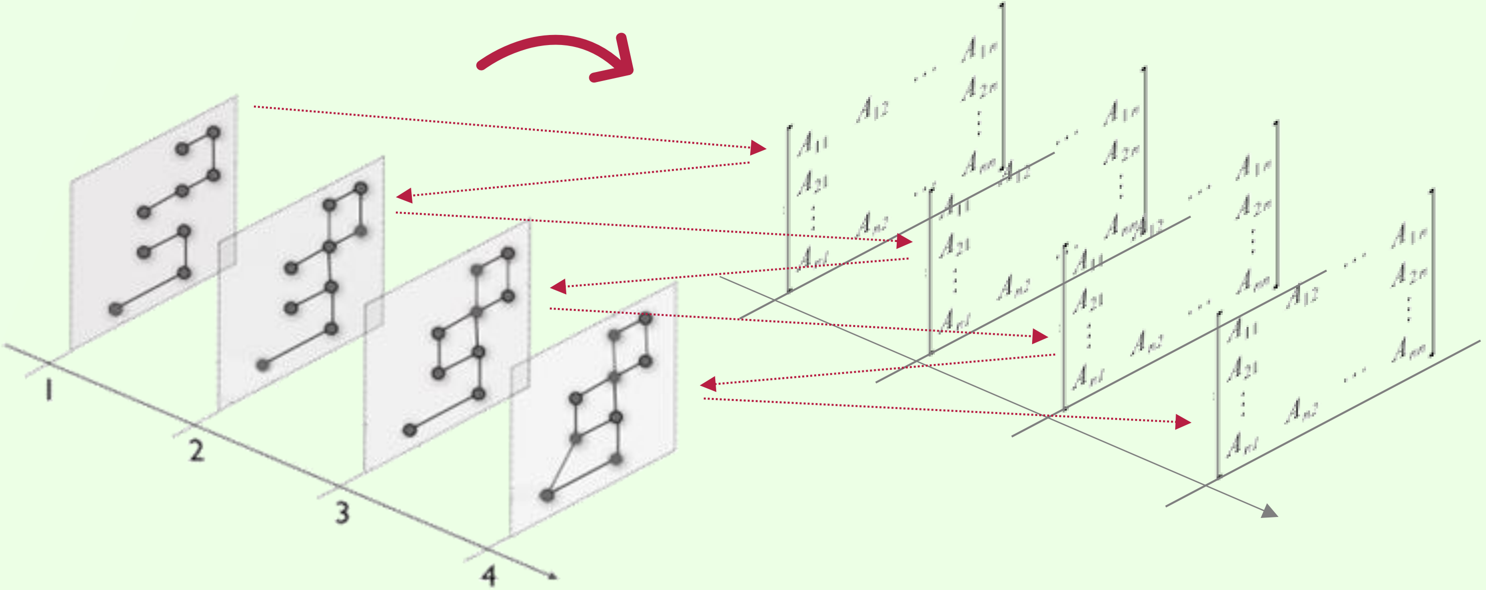
ECIR 2020 | ONLINE | 14-17 APRIL 2020

TEMPORAL LATENT SPACE MODELING

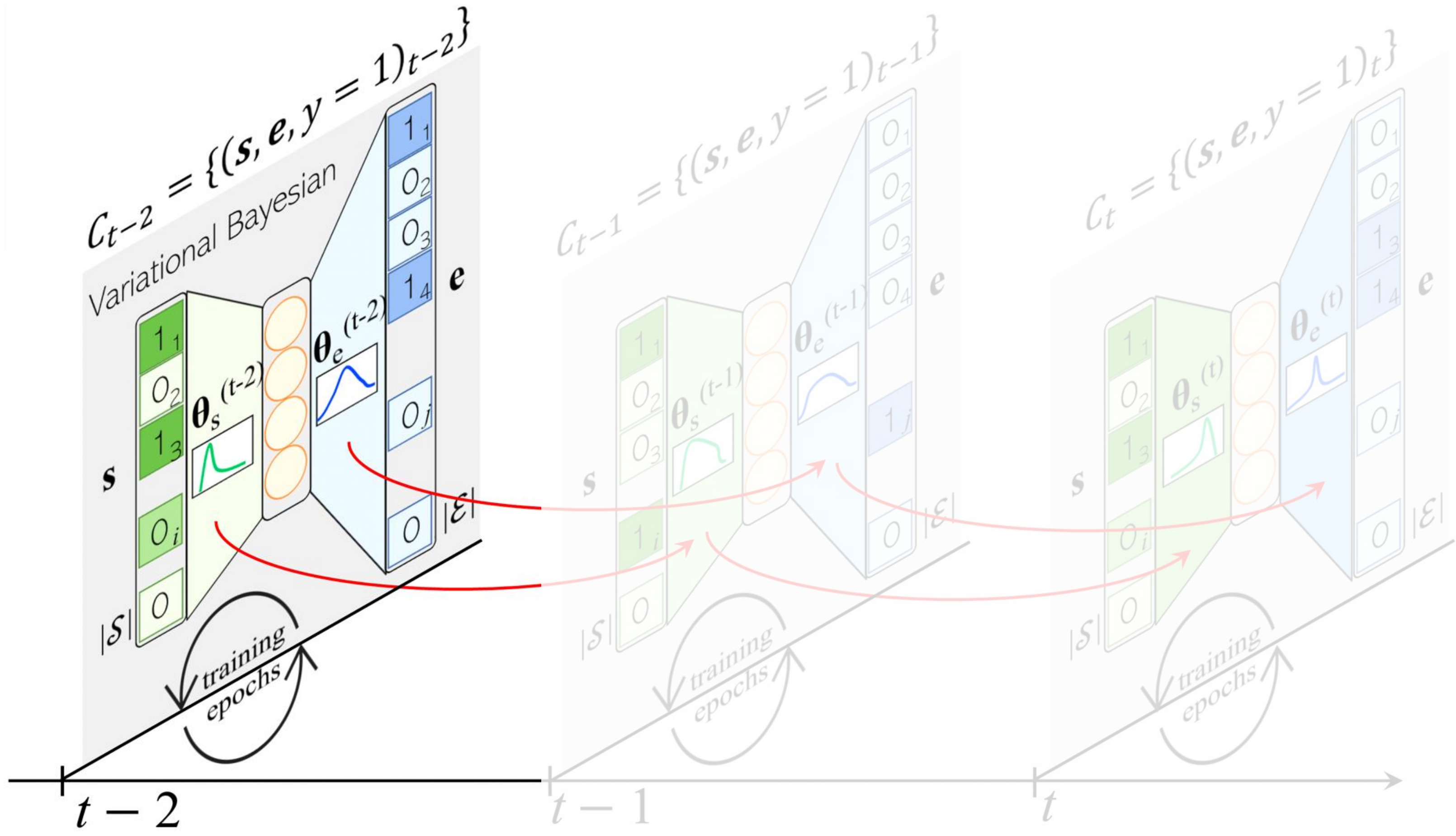
local Block Coordinate Gradient Descent (Zhu et al. TKDE 2016)

$$\arg\min \left[\sum_{t=1}^{T} \sum_{u,v \in \mathbb{U}} |w(u,v:t) - \mathbf{y}_{ut}\mathbf{y}_{vt}^{\top}|_F^2 + \lambda \sum_{t=1}^{T} \sum_{u \in \mathbb{U}} (1 - \mathbf{y}_{ut}\mathbf{y}_{u(t-1)}^{\top}) \right]$$

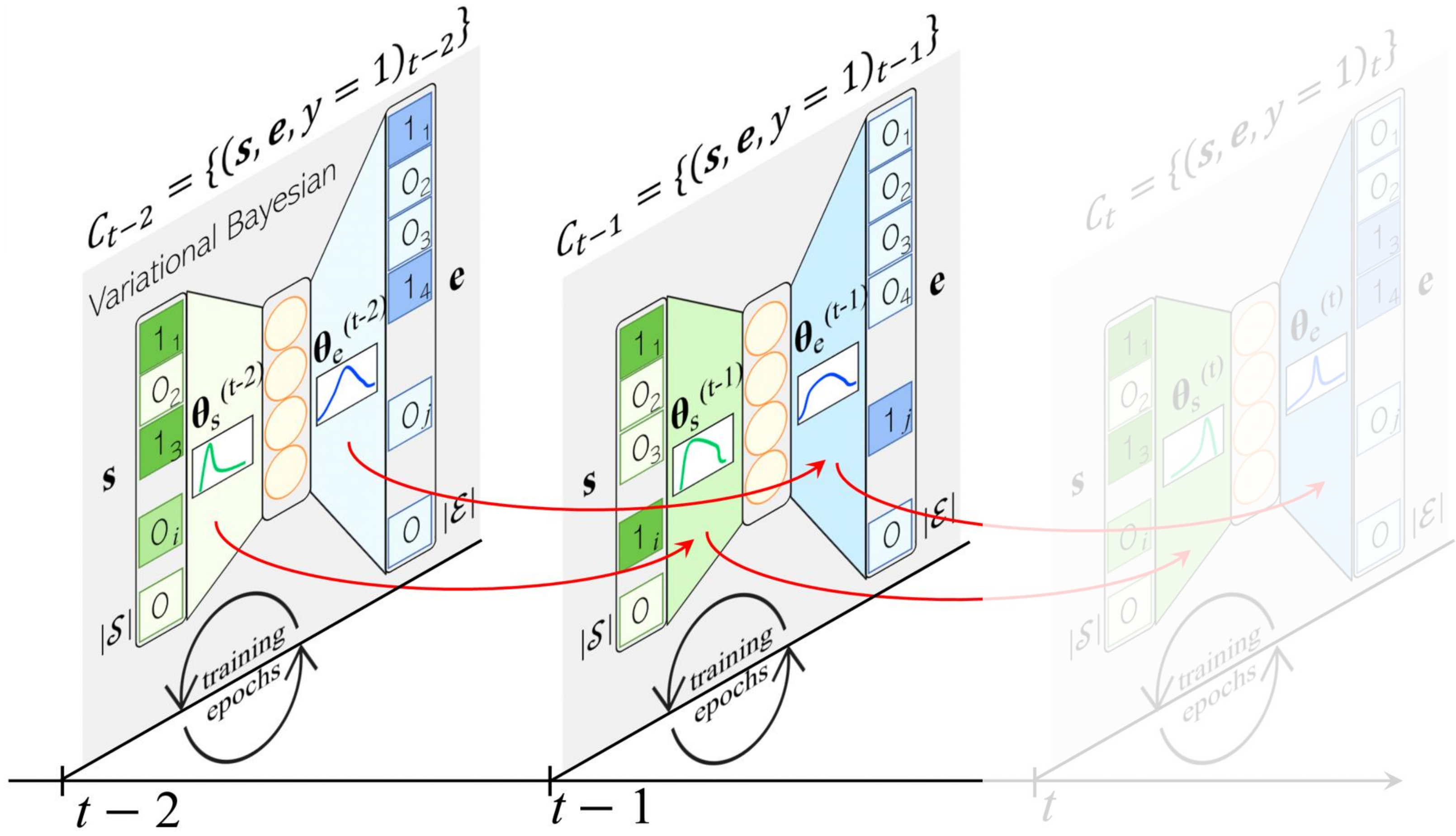
$$\forall u \in \mathbb{U}; \mathbf{y}_{ut} \geq 0, \mathbf{y}_{ut}\mathbf{y}_{ut}^{\top} = 1$$



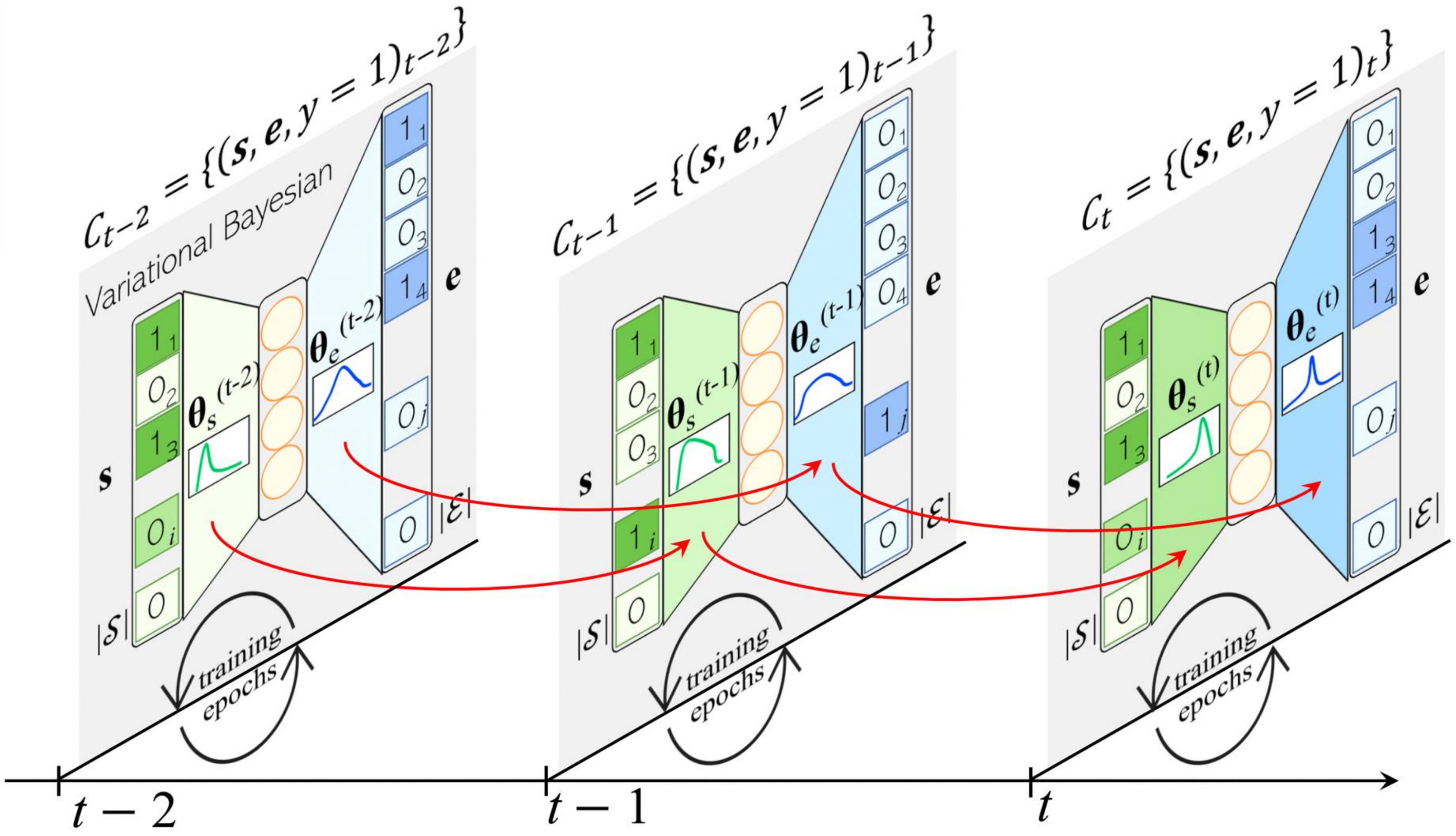
Fani et al. Temporal latent space modeling for community prediction. ECIR 2020



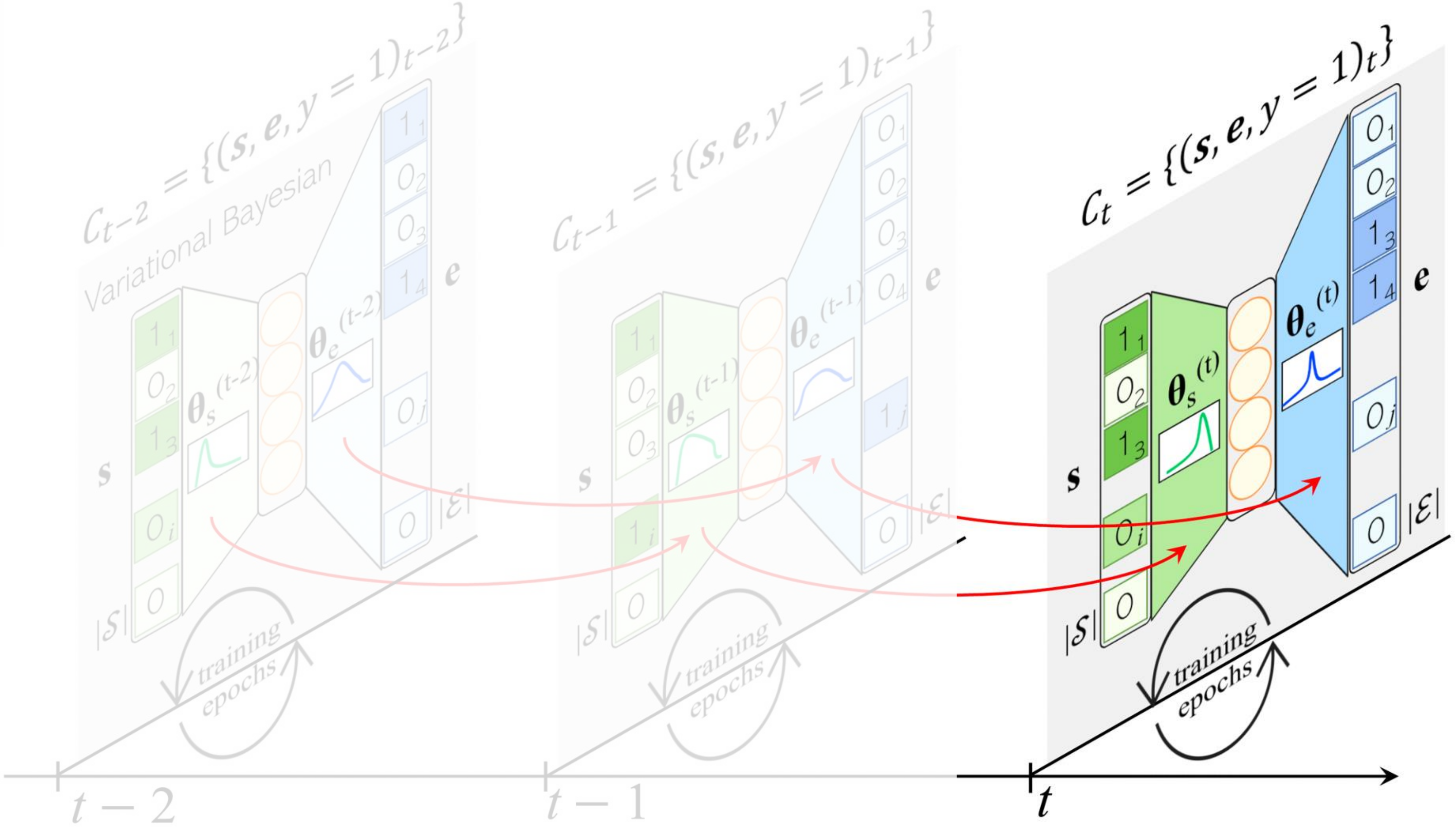
Streaming Training Strategy



Streaming Training Strategy



Streaming Training Strategy



T+1

Streaming Training Strategy

Table 1: Statistics of the raw and preprocessed datasets.

	dblp		uspt		imdb		gith	
	raw	filtered	raw	filtered	raw	filtered	raw	filtered
#teams	4,877,383	99,375	7,068,508	152,317	507,034	32,059	132,851	11,312
#unique experts	5,022,955	14,214	3,508,807	12,914	876,981	2,011	452,606	2,686
#unique skills	89,504	29,661	241,961	67,315	28	23	20	19
avg #expert per team	3.06	3.29	2.51	3.79	1.88	3.98	5.52	7.53
avg #skill per team	8.57	9.71	6.29	9.97	1.54	1.76	1.37	1.57
avg #team per expert	2.97	23.02	5.05	44.69	1.09	62.45	1.62	31.72
avg #skill per expert	16.73	96.72	19.49	102.53	1.59	10.85	2.03	5.18
#team w/ single expert	768,956	0	2,578,898	0	322,918	0	0	0
#team w/ single skill	5,569	56	939,955	8,110	315,503	15,180	69,131	6014
Timespan (raw)	1979 – 2018		1976 – 2019		1914 – 2020		2008 – 2022	

Table 1: Statistics of the raw and preprocessed datasets.

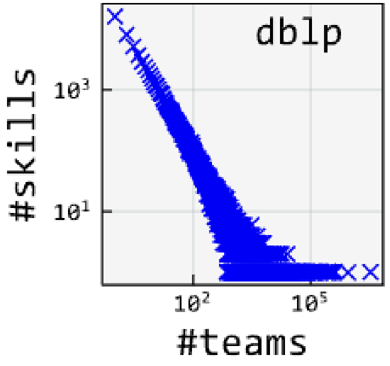
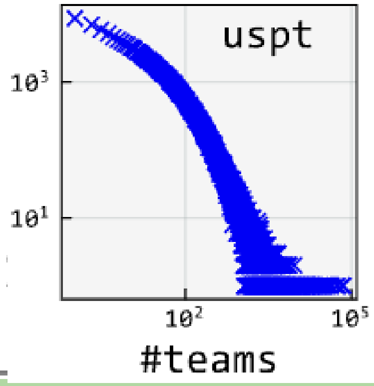
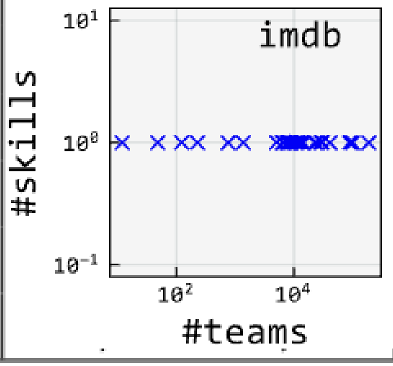
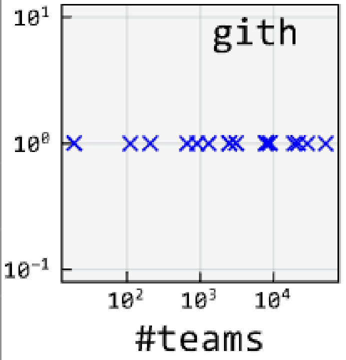
	dblp		uspt		imdb		gith	
	raw	filtered	raw	filtered	raw	filtered	raw	filtered
#teams	4,877,383	99,375	7,068,508	152,317	507,034	32,059	132,851	11,312
#unique experts	5,022,955	14,214	3,508,807	12,914	876,981	2,011	452,606	2,686
#unique skills	89,504	29,661	241,961	67,315	28	23	20	19
avg #expert per team								
avg #skill per team								
avg #team per expert								
avg #skill per expert								
#team w/ single expert								
#team w/ single skill								
Timespan (raw)	1979 – 2018		1976 – 2019		1914 – 2020		2008 – 2022	

Table 1: Statistics of the raw and preprocessed datasets.

	dblp		uspt		imdb		gith	
	raw	filtered	raw	filtered	raw	filtered	raw	filtered
#teams	4,877,383	99,375	7,068,508	152,317	507,034	32,059	132,851	11,312
#unique experts	5,022,955	14,214	3,508,807	12,914	876,981	2,011	452,606	2,686
#unique skills	89,504	29,661	241,961	67,315	28	23	20	19
avg #expert per team	3.06	3.29	2.51	3.79	1.88	3.98	5.52	7.53
avg #skill per team	8.57	9.71	6.29	9.97	1.54	1.76	1.37	1.57
avg #team per expert	2.97	23.02	5.05	44.69	1.09	62.45	1.62	31.72
avg #skill per expert	16.73	96.72	19.49	102.53	1.59	10.85	2.03	5.18
#team w/ single expert	768,956	0	2,578,898	0	322,918	0	0	0
#team w/ single skill	5,569	56	939,955	8,110	315,503	15,180	69,131	6014
Timespan (raw)	1979 – 2018		1976 – 2019		1914 – 2020		2008 – 2022	

RQ1: Does moving embeddings of experts and skill through time in the latent space improve the performance of neural models for the prediction of future successful teams?

RQ2: Does adding time explicitly to the input embeddings of skills boost neural models performance?

RQ3: Is the impact of our proposed training strategy consistent across datasets from various domains with distinct statistical distributions?

Low values of evaluation metrics for practical application

- Primarily due to the simplicity of the architectures
- Small number of epochs

Our main goal is not to report state-of-the-art results for a novel model

But to showcase the synergistic effects of streaming training strategy

Table 2: Average performance of 5-fold neural models on the test set.

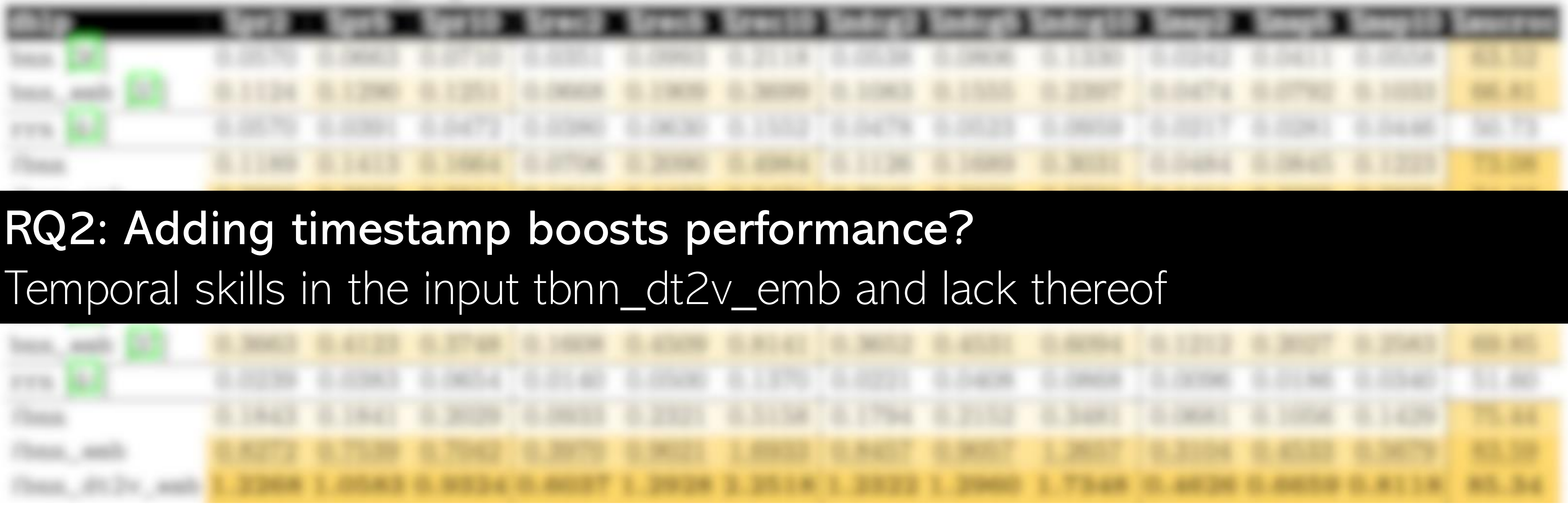
RQ1: Randomly Shuffled vs. Streaming?

Variational Bayesian neural network with streaming (tbnn-*) and lack thereof (bnn-*)

Table 2: Average performance of 5-fold neural models on the test set.

dblp	%pr2	%pr5	%pr10	%rec2	%rec5	%rec10	%ndcg2	%ndcg5	%ndcg10	%map2	%map5	%map10	%aucroc
bnn [36]	0.0570	0.0663	0.0710	0.0351	0.0993	0.2118	0.0538	0.0806	0.1330	0.0242	0.0411	0.0558	63.52
bnn_emb [35]	0.1124	0.1290	0.1251	0.0668	0.1909	0.3699	0.1083	0.1555	0.2397	0.0474	0.0792	0.1033	66.81
rrn [44]	0.0570	0.0391	0.0472	0.0380	0.0630	0.1552	0.0478	0.0523	0.0959	0.0217	0.0281	0.0446	50.73
*t*bnn	0.1189	0.1413	0.1664	0.0706	0.2090	0.4984	0.1126	0.1689	0.3031	0.0484	0.0845	0.1223	73.08
*t*bnn_emb	0.2996	0.2938	0.2811	0.1816	0.4433	0.8431	0.3048	0.3860	0.5721	0.1411	0.2095	0.2635	74.83
*t*bnn_dt2v_emb	**0.4299**	**0.3973**	**0.3612**	**0.2601**	**0.5963**	**1.0801**	**0.4284**	**0.5221**	**0.7465**	**0.1947**	**0.2864**	**0.3520**	**77.01**
uspt													
bnn [36]	0.0657	0.0769	0.0910	0.0353	0.0976	0.2212	0.0655	0.0883	0.1481	0.0266	0.0433	0.0592	64.54
bnn_emb [35]	0.3663	0.4123	0.3748	0.1608	0.4509	0.8141	0.3652	0.4531	0.6094	0.1212	0.2027	0.2583	69.85
rrn [44]	0.0239	0.0383	0.0654	0.0140	0.0500	0.1370	0.0221	0.0408	0.0868	0.0096	0.0186	0.0340	51.60
*t*bnn	0.1843	0.1841	0.2029	0.0933	0.2321	0.5158	0.1794	0.2152	0.3481	0.0681	0.1056	0.1429	75.44
*t*bnn_emb	0.8272	0.7539	0.7042	0.3970	0.9021	1.6933	0.8457	0.9057	1.2657	0.3104	0.4533	0.5679	83.59
*t*bnn_dt2v_emb	**1.2268**	**1.0583**	**0.9324**	**0.6037**	**1.2928**	**2.2518**	**1.2322**	**1.2960**	**1.7348**	**0.4626**	**0.6659**	**0.8118**	**85.34**

Table 2: Average performance of 5-fold neural models on the test set.



RQ2: Adding timestamp boosts performance?

Temporal skills in the input tbnn_dt2v_emb and lack thereof

Table 2: Average performance of 5-fold neural models on the test set.

dblp	%pr2	%pr5	%pr10	%rec2	%rec5	%rec10	%ndcg2	%ndcg5	%ndcg10	%map2	%map5	%map10	%aucroc
bnn [36]	0.0570	0.0663	0.0710	0.0351	0.0993	0.2118	0.0538	0.0806	0.1330	0.0242	0.0411	0.0558	63.52
bnn_emb [35]	0.1124	0.1290	0.1251	0.0668	0.1909	0.3699	0.1083	0.1555	0.2397	0.0474	0.0792	0.1033	66.81
rrn [44]	0.0570	0.0391	0.0472	0.0380	0.0630	0.1552	0.0478	0.0523	0.0959	0.0217	0.0281	0.0446	50.73
*t*bnn	0.1189	0.1413	0.1664	0.0706	0.2090	0.4984	0.1126	0.1689	0.3031	0.0484	0.0845	0.1223	73.08
*t*bnn_emb	0.2996	0.2938	0.2811	0.1816	0.4433	0.8431	0.3048	0.3860	0.5721	0.1411	0.2095	0.2635	74.83
*t*bnn_dt2v_emb	**0.4299**	**0.3973**	**0.3612**	**0.2601**	**0.5963**	**1.0801**	**0.4284**	**0.5221**	**0.7465**	**0.1947**	**0.2864**	**0.3520**	**77.01**
uspt													
bnn [36]	0.0657	0.0769	0.0910	0.0353	0.0976	0.2212	0.0655	0.0883	0.1481	0.0266	0.0433	0.0592	64.54
bnn_emb [35]	0.3663	0.4123	0.3748	0.1608	0.4509	0.8141	0.3652	0.4531	0.6094	0.1212	0.2027	0.2583	69.85
rrn [44]	0.0239	0.0383	0.0654	0.0140	0.0500	0.1370	0.0221	0.0408	0.0868	0.0096	0.0186	0.0340	51.60
*t*bnn	0.1843	0.1841	0.2029	0.0933	0.2321	0.5158	0.1794	0.2152	0.3481	0.0681	0.1056	0.1429	75.44
*t*bnn_emb	0.8272	0.7539	0.7042	0.3970	0.9021	1.6933	0.8457	0.9057	1.2657	0.3104	0.4533	0.5679	83.59
*t*bnn_dt2v_emb	**1.2268**	**1.0583**	**0.9324**	**0.6037**	**1.2928**	**2.2518**	**1.2322**	**1.2960**	**1.7348**	**0.4626**	**0.6659**	**0.8118**	**85.34**

Table 2: Average performance of 5-fold neural models on the test set.

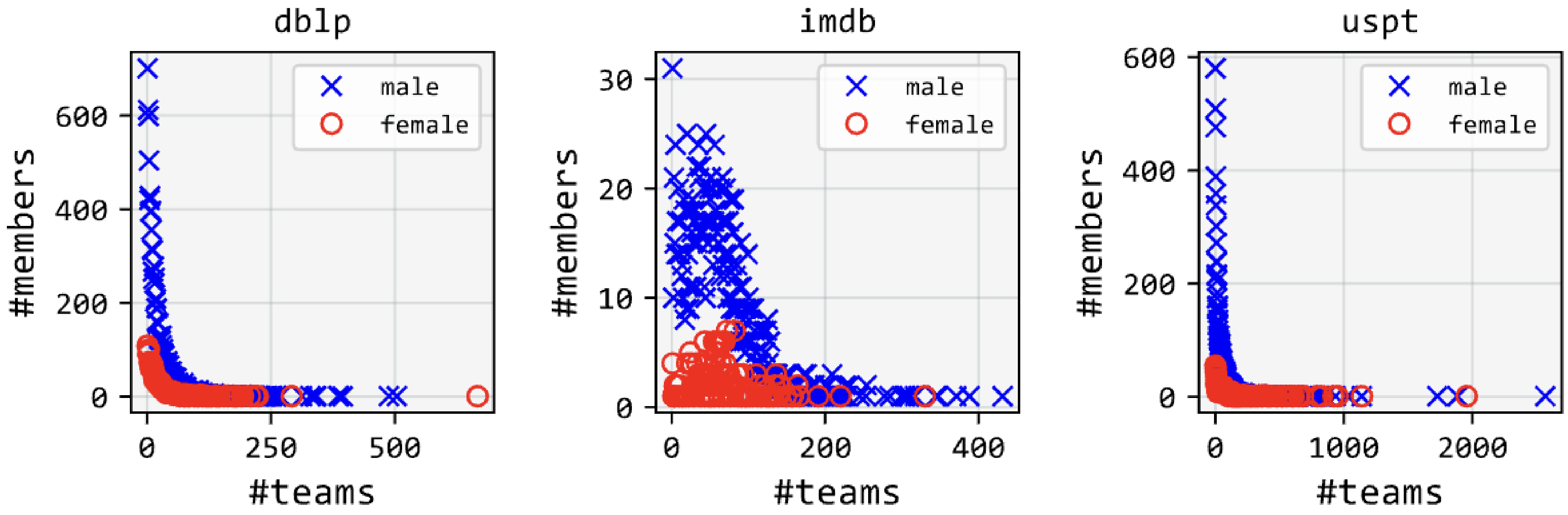
RQ3: Is the impact consistent across datasets?

dblp, uspt vs. imdb, gith

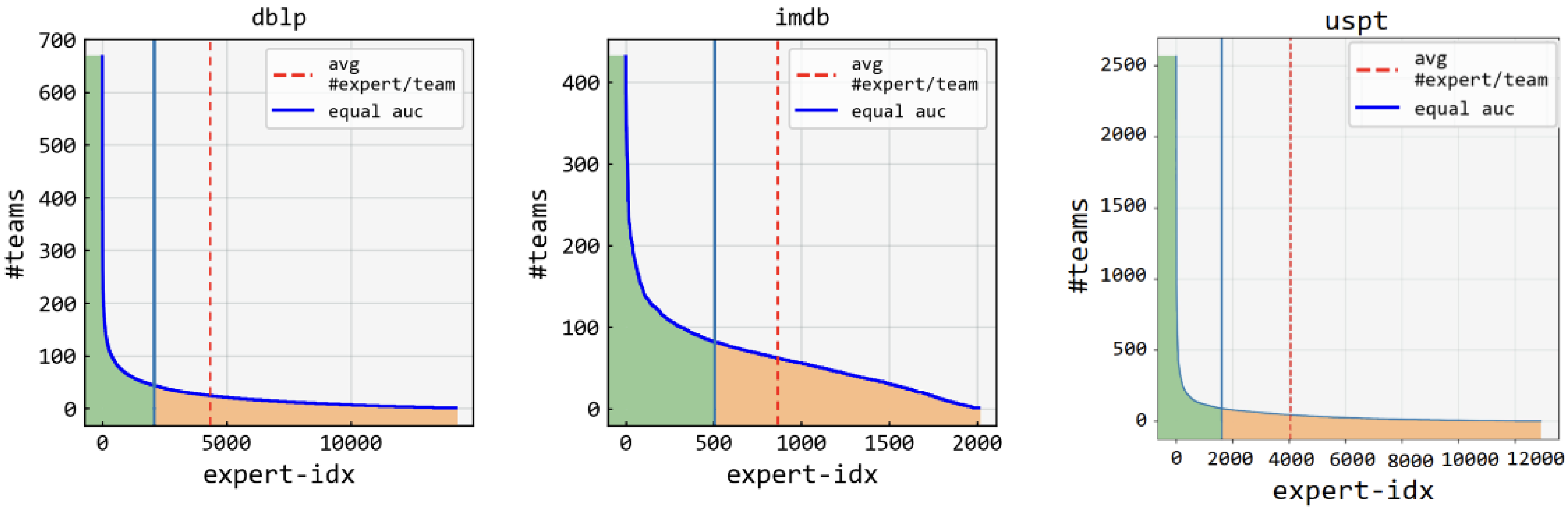
imdb													
bnn [36]	0.2128	0.5106	0.4255	0.1418	0.8511	1.3050	0.1646	0.5699	0.7848	0.0709	0.2600	0.3148	51.16
bnn_emb [35]	0.4255	0.5106	0.6383	0.2837	0.8511	1.9574	0.3292	0.5923	1.1358	0.1418	0.2813	0.4389	51.82
rrn [44]	0.0000	0.8511	0.8511	0.0000	1.4184	2.8369	0.0000	0.8163	1.4606	0.0000	0.3191	0.6265	52.22
*t*bnn	0.8511	**1.5319**	1.4043	0.5319	**2.4610**	4.4965	0.7548	1.7381	2.6829	0.3369	0.8215	1.1674	63.46
*t*bnn_emb	0.8511	1.1064	1.0638	0.5674	1.7518	1.3262	0.9474	1.4848	2.2007	0.4965	0.8138	1.0099	**66.87**
*t*bnn_dt2v_emb	**1.9149**	1.1915	**1.4468**	**1.2411**	1.9504	**4.5532**	**1.8667**	**1.8703**	**3.0303**	**0.9043**	**1.1099**	**1.4293**	66.56
gith													
bnn [36]	3.0693	2.8515	2.6931	1.2164	2.8846	5.1174	3.1365	3.2893	4.2340	1.0104	1.5706	2.1633	56.18
bnn_emb [35]	**7.3267**	**4.7129**	3.3861	**3.5441**	**5.1580**	**6.1885**	**6.4753**	**5.8418**	**6.2665**	**2.3424**	**3.0822**	**3.3837**	62.65
rrn [44]	0.0000	0.1980	0.0990	0.0000	0.0619	0.0619	0.0000	0.1679	0.1090	0.0000	0.0206	0.0206	52.26
*t*bnn	3.8614	2.8515	2.3564	1.8801	3.1525	4.5754	4.3319	3.9721	4.5031	1.8025	2.3978	2.8768	56.65
*t*bnn_emb	4.9505	3.5248	3.1287	1.9434	3.0770	4.3718	5.0849	4.4715	4.9844	1.6957	2.1431	2.5949	62.20
*t*bnn_dt2v_emb	5.7426	4.5941	**3.8020**	2.1874	3.8474	4.7855	5.6081	5.3287	5.6670	1.7131	2.4258	2.7858	**64.89**

Gihub: 2008 – 2022


Future Direction: Bias



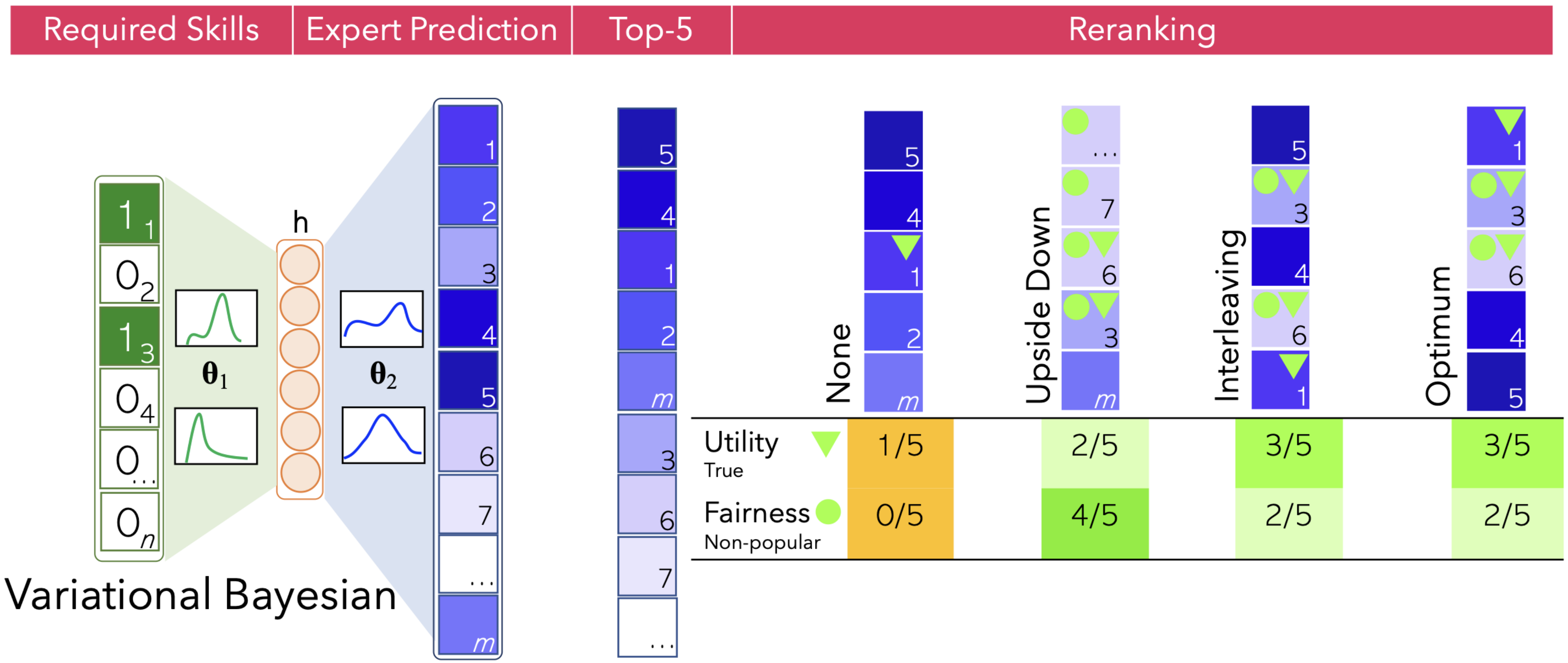
	dblp	imdb	uspt
%female experts	14.20%	12.30%	13.80%



	dblp	imdb	uspt
%popular experts (avg)	31.30%	42.60%	31.40%

Adila*: Fairness-Aware Team Formation

* عادلة, feminine Arabic given name, meaning just and fair



Post-processing Fair Reranking

github.com/fani-lab/OpeNTF

fani-lab / OpeNTF

Type / to search

Code | Issues 58 | Pull requests | Discussions | Actions | Projects | Wiki | Security 656 | Insights | Settings

OpeNTF Public

Edit Pins | Unwatch 4 | Fork 10 | Star 15

Publish your first package

Contributors 13

Deployments 27

github-pages last month

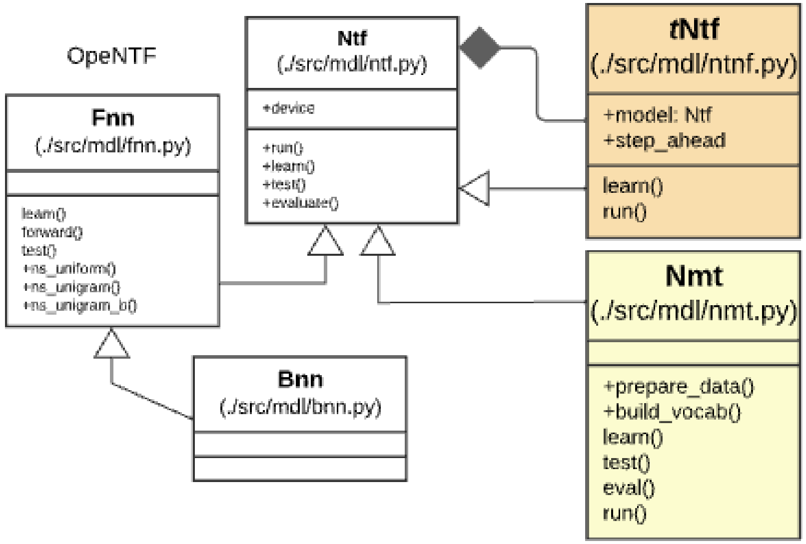
+ 26 deployments

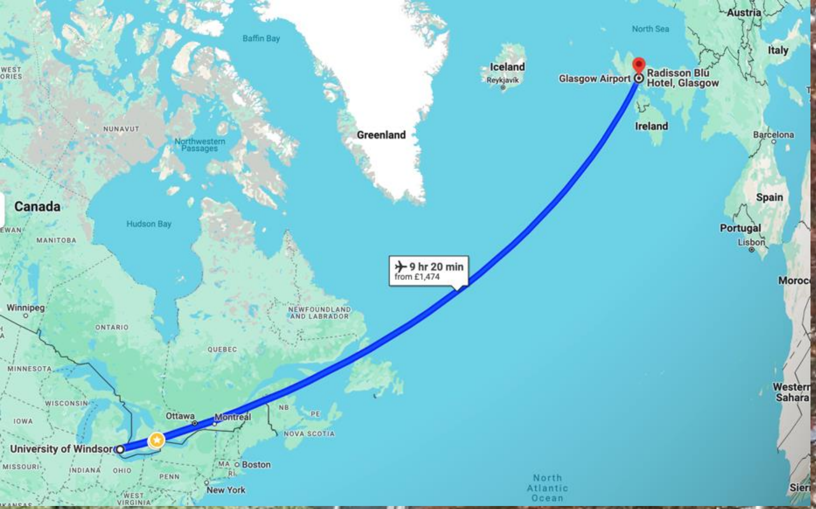
README | License

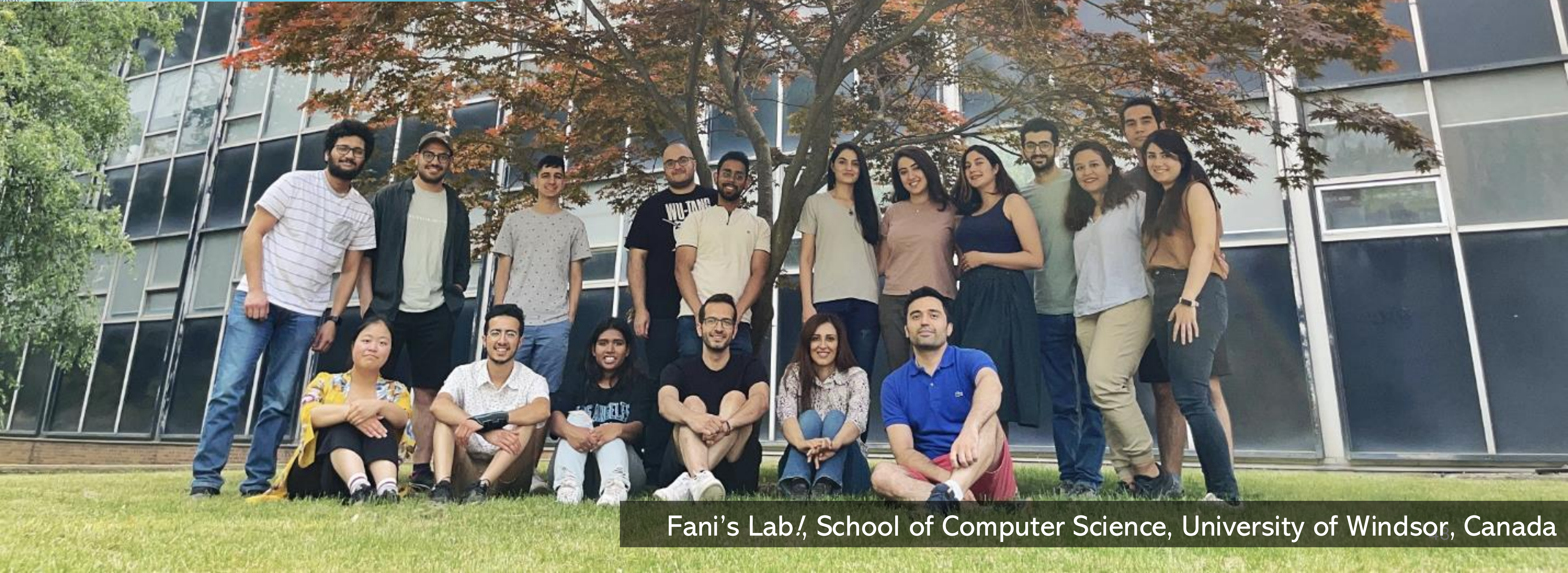
OpeNTF: An Open-Source Neural Team Formation Benchmark Library

Team formation involves selecting a team of skillful experts who will, more likely than not, accomplish a task. Researchers have proposed a rich body of computational methods to automate the traditionally tedious and error-prone manual process. We previously released OpeNTF, an open-source framework hosting canonical neural models as the cutting-edge class of approaches, along with large-scale training datasets from varying domains. In this paper, we contribute OpeNTF2 that extends the initial release in two prime directions. (1) The first of its kind in

- Fairness aware Team Formation
- Datasets and Parallel Preprocessing
- Non-Temporal Neural Team Formation
- Temporal Neural Team Prediction
- Model Architecture
- Negative Sampling Strategies







Fani's Lab!, School of Computer Science, University of Windsor, Canada


Reza
Arman
Mahdis
Hossein

UMAP '24
CAGLIARI·ITALY

Home Key Dates Calls Program Attend Organizers Sponsors

Collaborative Team Recommendation for Skilled Users:

Objectives, Techniques, and New perspectives


A slide for people affected by the disaster of wars ...
Embrace - Sculpted by Hans Schleeh, born in Germany and emigrated to Canada in 1951