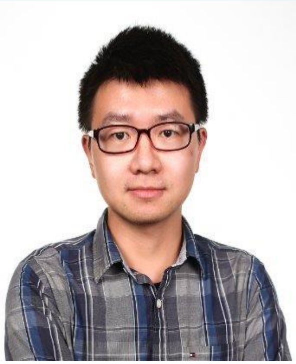
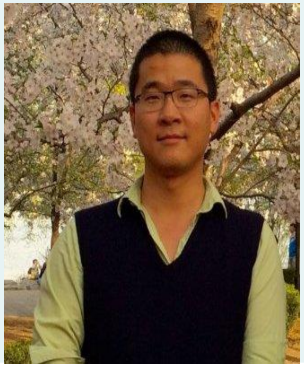
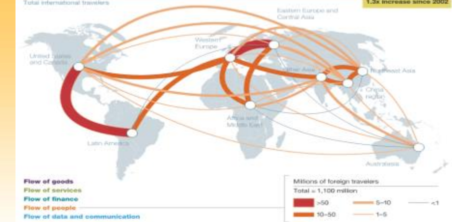
NEMO: Next Career Move Prediction with Contextual Embedding

Presenter: **Jaewon Yang**

Liangyue Li (ASU)	How Jing (LinkedIn)	Hanghang Tong (ASU)	Jaewon Yang (LinkedIn)	Qi He (LinkedIn)	Bee-Chung Chen (LinkedIn)
					

Labor Flow



- Labor flow: How people change jobs
- *Why do we care?*
  - Match supply with demand
  - Circulate knowledge and drive innovation
- Large-scale studies possible with web (e.g., LinkedIn)
- We focus on predicting individual's career moves on large scale

State of the Arts

- **Macro-level**
  - Employer-to-employer flows[Bjelland+2011]
  - Labor flow networks[Guerrero+2013]
- **Micro-level**
  - How scientists move and transform careers [Deville+2014]
- **Job Recommendation**
  - Model matching users and job postings [Paparrizos+2011]

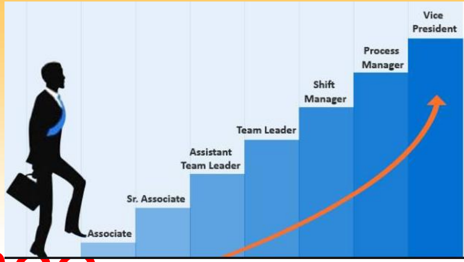
Next Career Move Prediction

- **Given**: the sequence of positions (title / company) of users up to time *T*, users' skill sets, education and locations
- **Predict**: a user's next position (title/company) after *T*

Design Objectives

- **Goal-1**: *using profile attributes*
  - Career moves reflect profile info (*skills, education, location*)
  - Helps for people with short career history
- **Challenges**
  - Single-valued (*final education*), Multi-valued (*skill set*)

Design Objectives



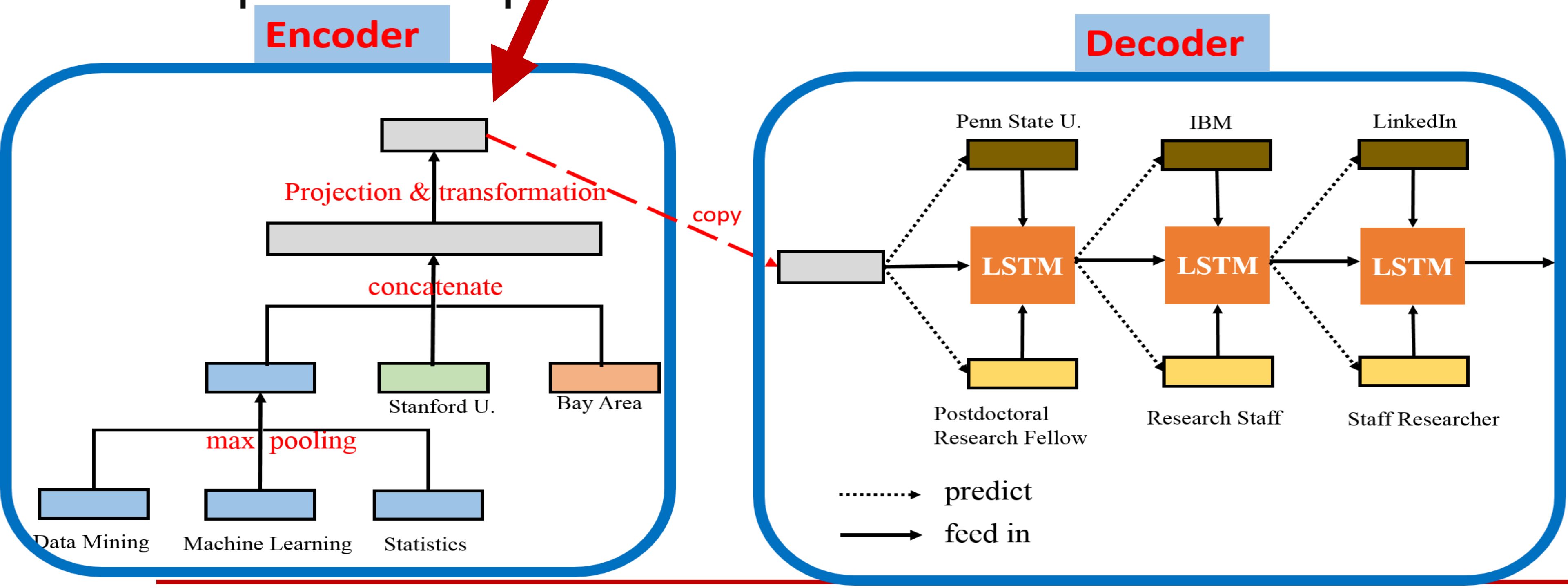
- **Goal-2**: *modeling position sequence*
  - Career moves reflect the history of one's past career path
  - Rare to switch to an entirely new field
- **Challenges**
  - Not enough to consider only current position
  - Model the entire sequence of job positions

Roadmap

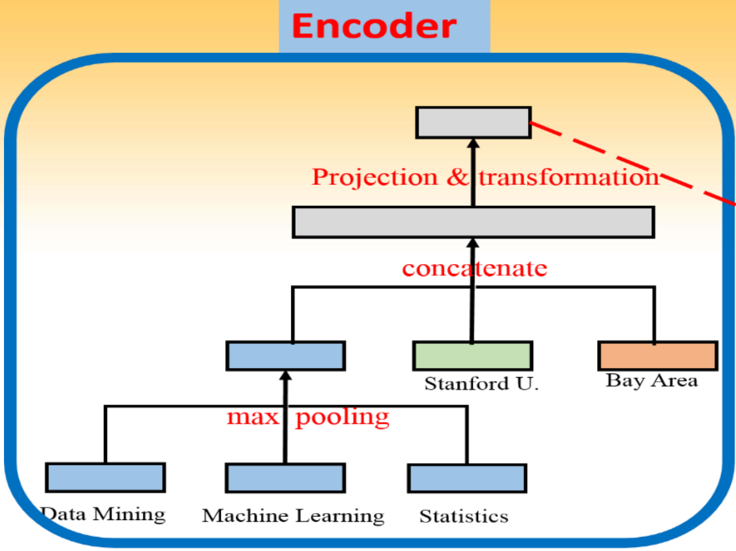
- Motivations
- **Proposed Solutions – NEMO**
- Empirical Evaluations
- Conclusions

An Encoder-Decoder Architecture

- **Encoder**: maps profile contexts to a fixed-length vector (*profile context vector*)
- **Decoder**: maps *profile context vector* to sequence of positions



Encoding the profile contexts



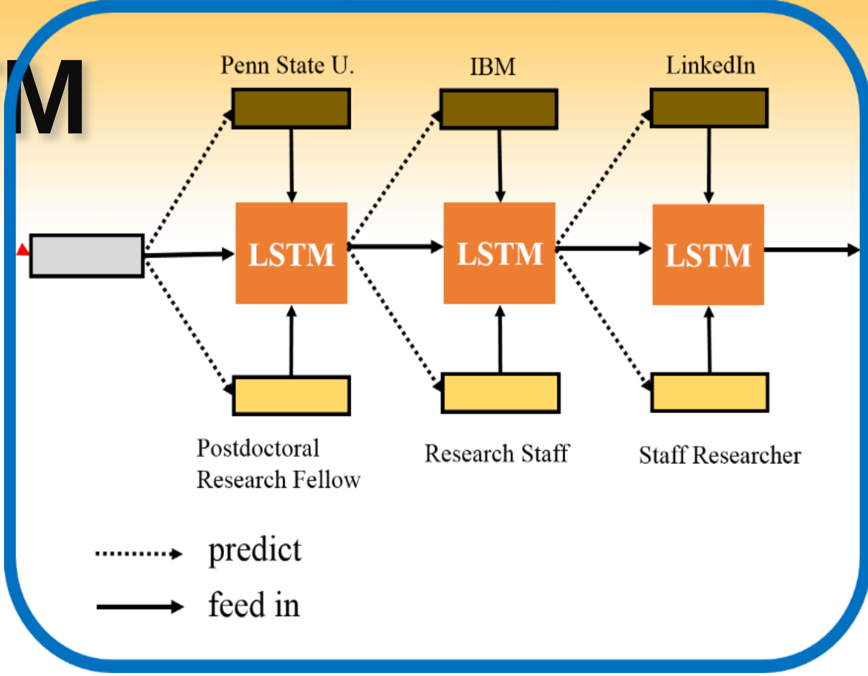
- Learning **embeddings** for skills s, schools h and locations r
- Aggregate all the profile contexts
  - Max-pooling on all the skills

$$\mathbf{s}^u = \max(\mathbf{s}_1^u, \mathbf{s}_2^u, \ldots, \mathbf{s}_m^u)$$

  - Concatenate skill with school/location, feed to a one-layer neural network

$$\mathbf{v}_u = \tanh\left(\mathbf{W}_v[\mathbf{s}^u, \mathbf{h}^u, \mathbf{r}^u]^T + \mathbf{b}_v\right)$$

Decoding with LSTM



- Long short-tem Memory Network (LSTM)
- Latent state vector
  - Initialized by the profile context vector
  - generates position (company / title), updated by LSTM
- Company/title independent given latent state

Learning and Prediction

- **Learning** - Maximizing the log probability of the observed career path
  - *Individual User's Career Path*

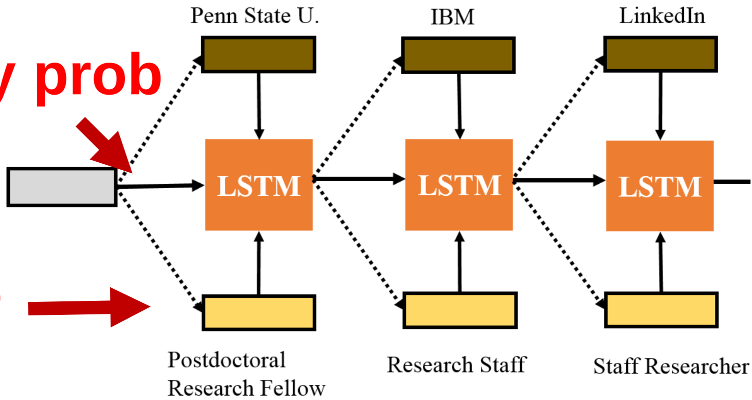
$$\log p(\mathcal{J}^u|\mathcal{S}^u, h^u, r^u) = \sum_{t=1}^{T} \log p(J_t^u|\mathcal{S}^u, h^u, r^u)$$

$$= \sum_{t=1}^{T} \Big[\log p_t(c_t^u|\mathcal{S}^u, h^u, r^u, c^u_{t'<t}, l^u_{t'<t})$$

$$+ \log p_t(l_t^u|\mathcal{S}^u, h^u, r^u, c^u_{t'<t}, l^u_{t'<t})\Big]$$

Company prob

Title prob



- **Prediction** – Given state vector, predict title, predict company

Roadmap

- Motivations
- Proposed Solutions – NEMO
- **Empirical Evaluations**
- Conclusions

Datasets

- Construct two datasets from LinkedIn
  - *Computer*: comp. software, internet, IT, etc
  - *Finance*: banking, investment, etc
- **Task**: predict next company/title after 12/01/2015, given position sequence up to 09/01/2015

Mean Percentile Rank Comparison

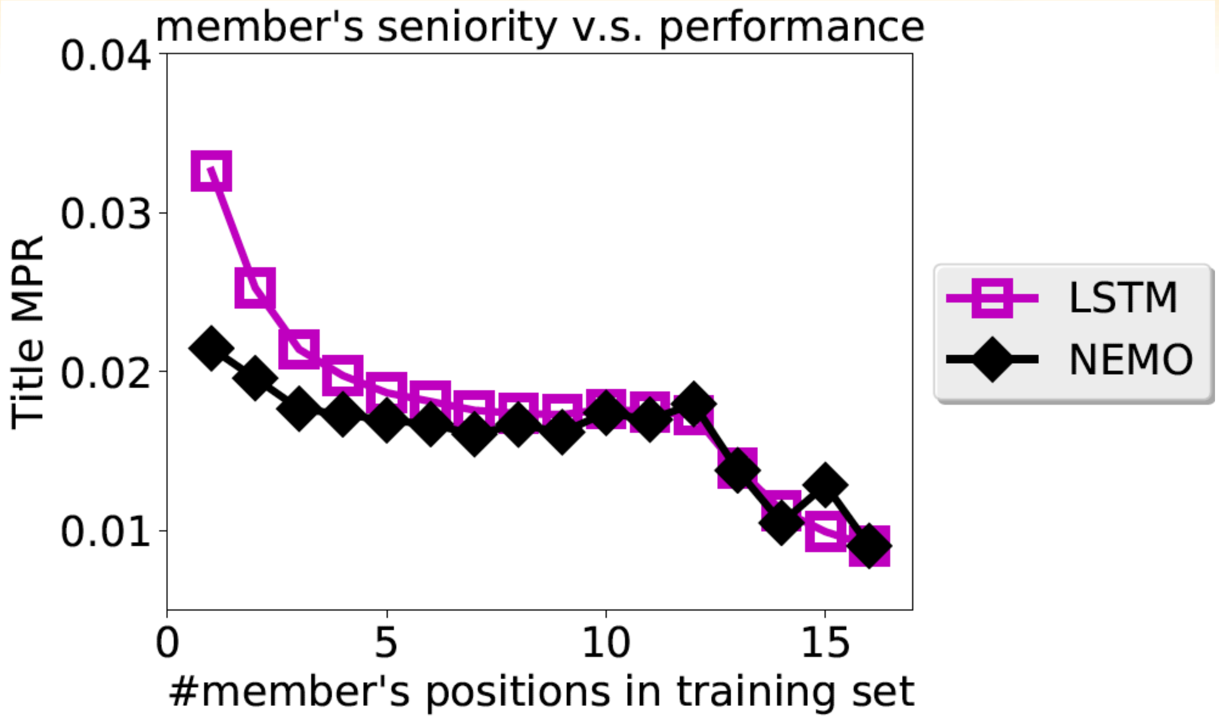
Smaller is better

	Computer		Finance	
	Company	Title	Company	Title
Profile Context	0.0512	0.0286	0.0403	0.0391
MC (no profile, last position only)	0.0542	0.0277	0.0496	0.0351
LSTM (No profile)	0.0432	0.0225	0.0411	0.0299
NEMO	**0.0299**	**0.0182**	**0.0260**	**0.0253**

LSTM> MC: entire sequence > last pos.

NEMO is best – profile context + career path

Stratification by Seniority



- NEMO > LSTM (x < 5): Profile context helps for juniors

Case Study

Previous position	Ground-truth next position	Top 5 predicted Company	Top 5 predicted Title
Senior project manager @Fidelity Investments Worked at airline before	Project Manager @Southwest Airlines	Fidelity Investments American Airlines Southwest Airlines Epsilon Bank of America	Senior Project Manager Project Manager Technical Project Manager Senior Technical Project Manager Program Manager
Software Architect/Tech Lead @ Bureau of Labor Statistics Worked at USPTO before	Consultant @ United States Patent and Trademark Office (USPTO)	Fannie Mae USPTO FINRA Lockheed Martin Freddie Mac	Technical Lead Senior Software Engineer Consultant Senior Consultant Solutions Architect

Sampling Career Path

User 1
SF Bay Area
CMU graduate
Skills:ML, DM, AI, Algorithms
1st job: ML Engineer@Google

→ ML engineer *@WhatsApp* → ML engineer *@Uber* → Data Scientist *@Facebook* → Engineer Lead *@LinkedIn*

Both have a rising career trajectory

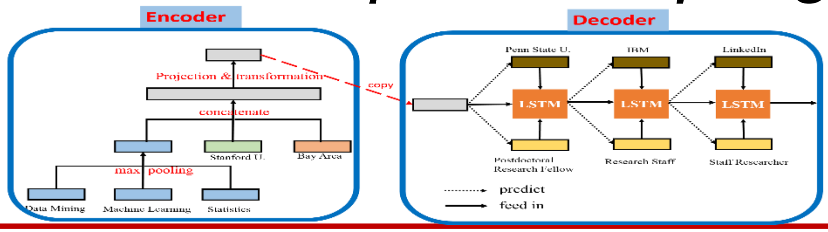
Cold-start Case

User 2
Greater New York Area
Harvard Business School
Skills:Financial services, investments

→ Investment banker *@Citi* → Technology Strategist *@Citi* → Relationship Manager *@Citi* → VP *@Morgan Stanley* → VP brokerage *@JPMorgan Chase*

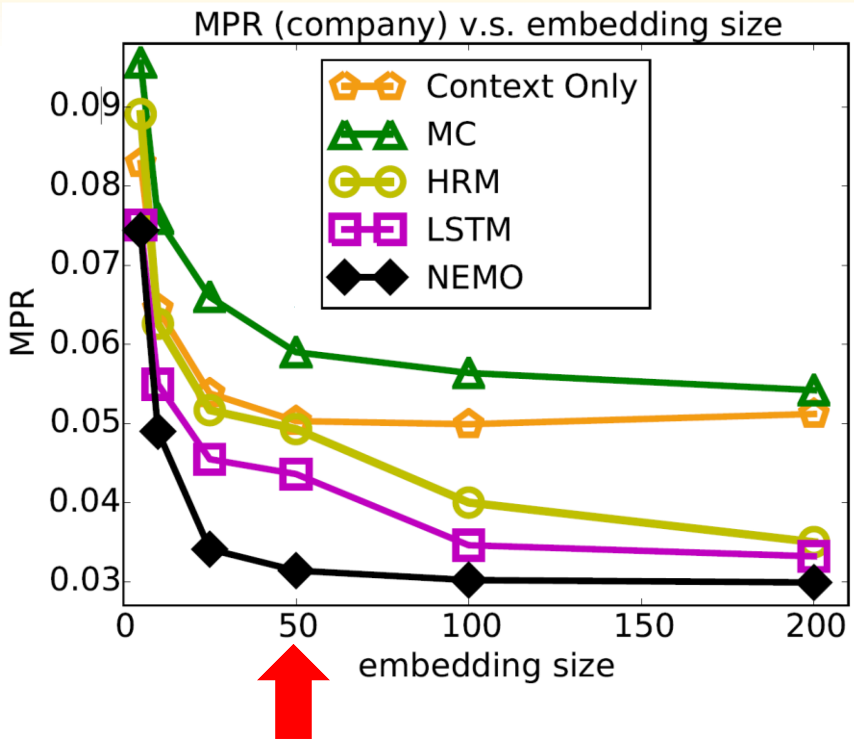
Conclusions

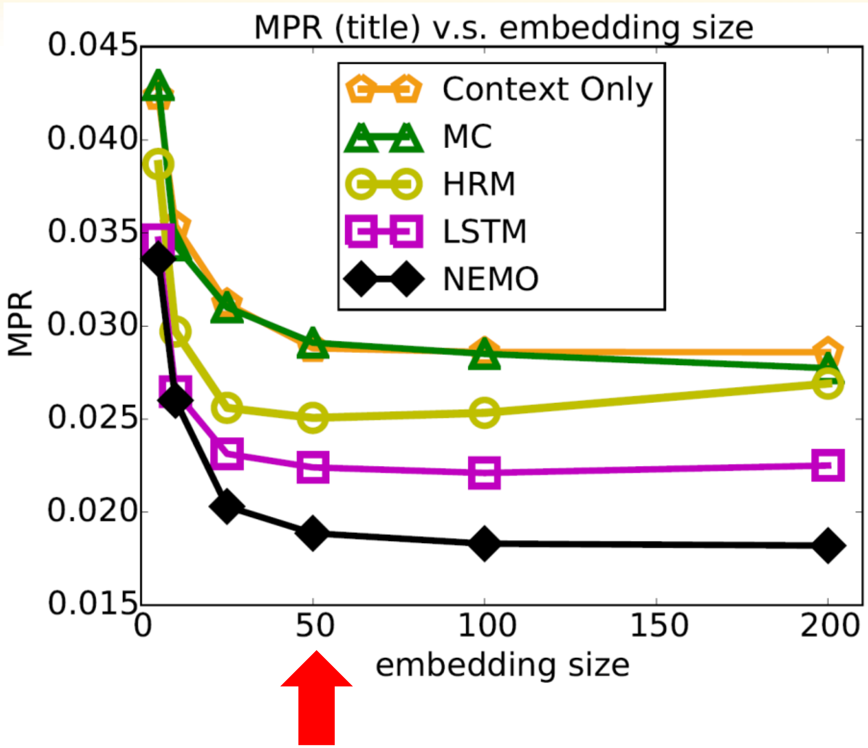
- **Problem Def**: next career move prediction
- **Design Objectives**: *profile context* + *position sequence*
- **Solutions**: **NEMO** – contextual LSTM
- **Results**:
  – Profile + Position sequence = better prediction
  – Insights from *career path sampling*



Thank you!

Varying Embedding Dimension

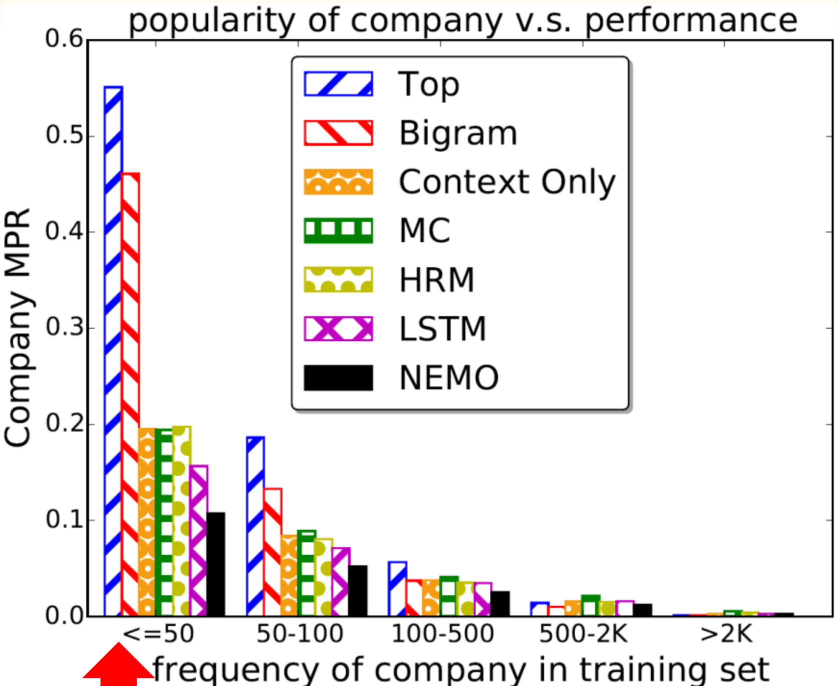


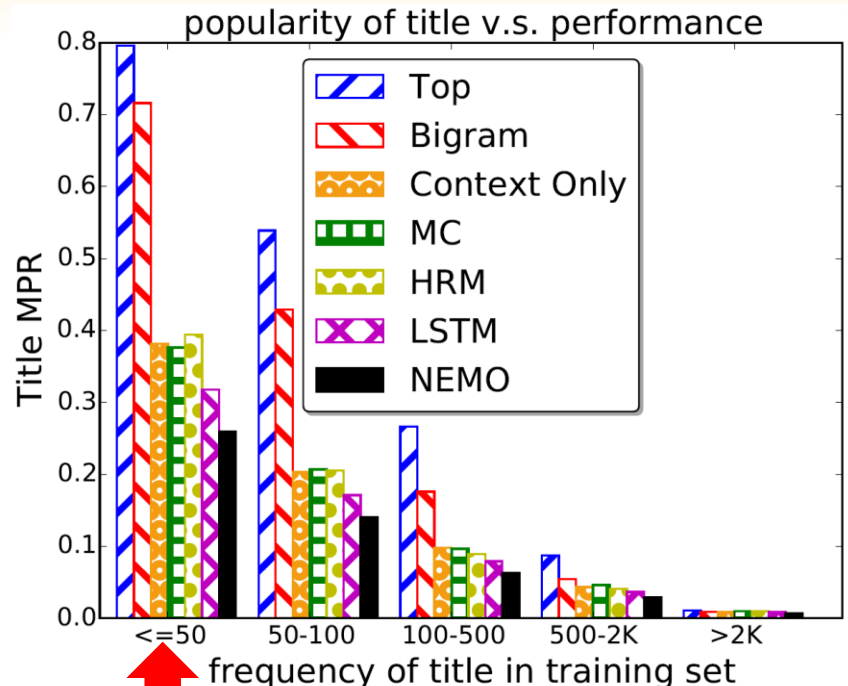


Diminishing return in all the models

Arizona State University

Stratification by Popularity





NEMO performs significantly better when company/title is rare