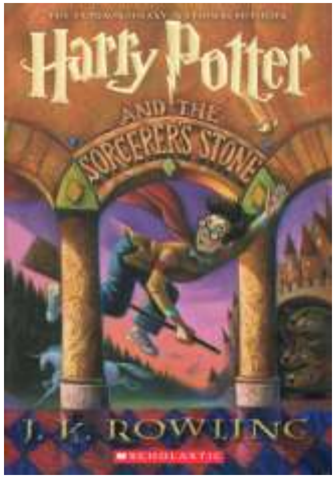
Predictability of popularity: Gaps between prediction and understanding

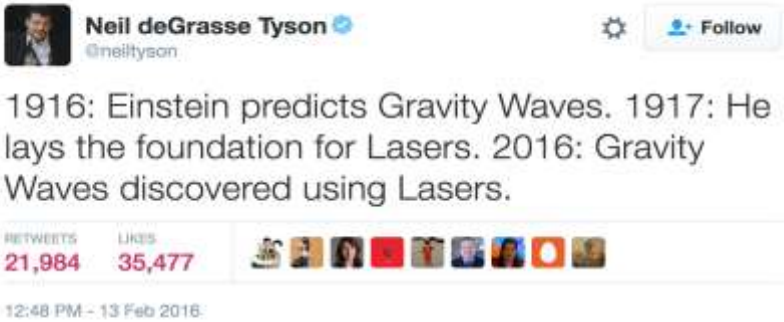
BENJAMIN SHULMAN, YELP

AMIT SHARMA, MICROSOFT RESEARCH

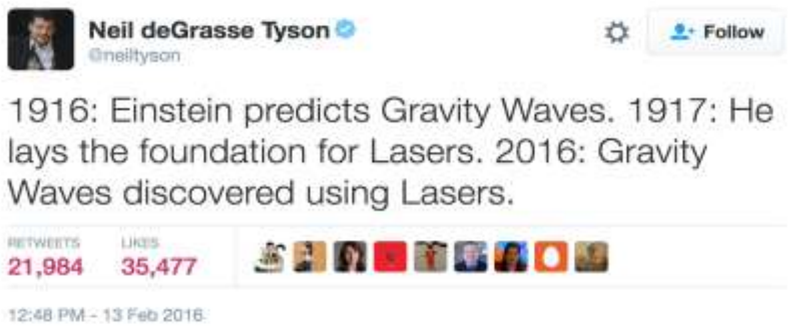
DAN COSLEY, CORNELL UNIVERSITY

Will it get popular?





Can we predict popularity?



Even with state-of-the-art features and prediction models, hard to predict *apriori* [Bakshy et al. 2011, Martin et al. 2016]

What if we could peek into early activity?

Szabo and Huberman (2010)

Pinto et al. (2012)

Zhao et al. (2015)

Cheng et al. (2014)

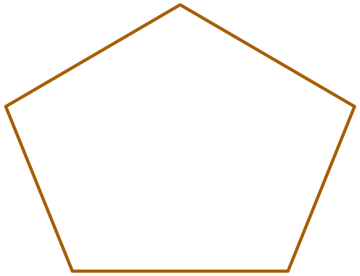
Possible to predict future popularity with high accuracy.

Prediction $\neq$ Understanding how items gain popularity

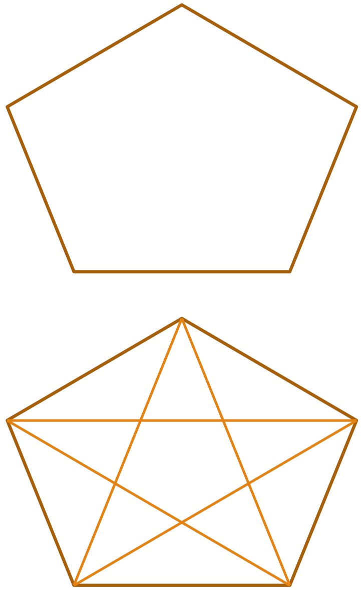
Varying conclusions from prior studies on predicting popularity

FEATURES	Content	Structural	Early Adopters	Temporal
Szabo and Huberman (2010)	**-**	**N**	**-**	**Y**
Tsur et al. (2012)	**Y**	**Y**	**-**	**Y**
Pinto et al. (2013)	**-**	**-**	**-**	**Y**
Yu et al. (2015)	**-**	**N**	**-**	**Y**
Romero et al. (2013)	**-**	**Y**	**-**	**-**
Cheng et al. (2014)	**-**	**Y**	**Y**	**Y**
Lerman et al. (2008)	**-**	**Y**	**Y**	**-**
Weng et al. (2013)	**-**	**Y**	**N**	**-**

Example: Does higher network density lead to higher popularity?



Lerman and Galstyan, 2008. Lower network density among early adopters lead to higher final popularity.



Romero et al., 2013. Both higher and lower density could lead to higher final popularity.

A multi-platform study on accuracy of popularity prediction and usefulness of different features.

Four datasets from different domains

LOVE SONG

437k users
5.8M songs
44M adoptions

SHARE URLs

737k users
58k tweets
2.7M adoptions

RATE BOOKS

183k users
10.9M tweets
28M adoptions

FAVORITE PHOTOS

252k users
1.3M tweets
33M adoptions

Given a set of items and data about their early adoptions, which among them are more likely to become popular?

- Use data about the first *K=5* early adoptions to predict whether final popularity at *T=28* is greater than median.

[Cheng et al. 2014]

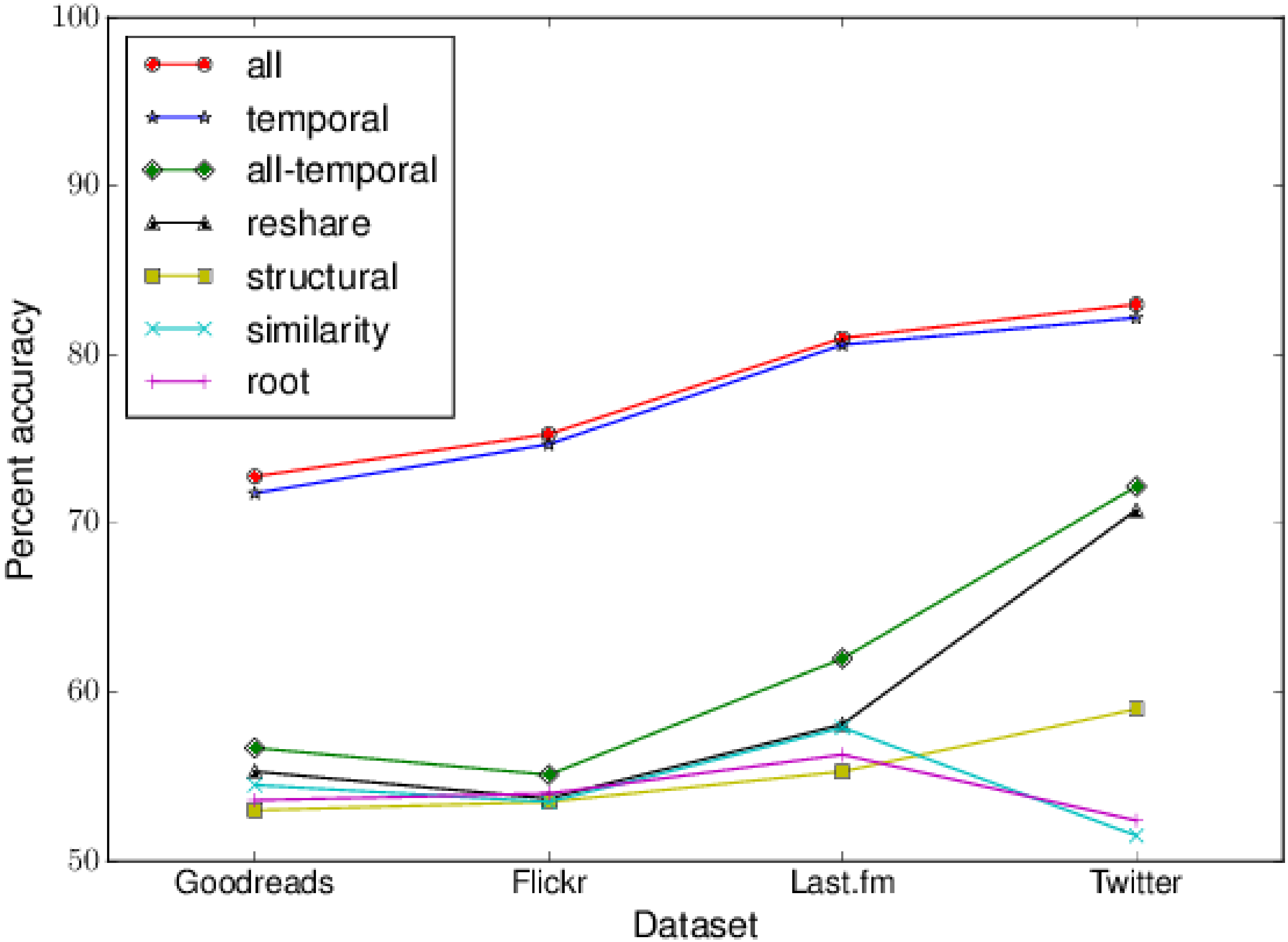
Features

Temporal: time for each adoption, time between adoptions

Structural: degree of early adopters, network density among early adopters

Early adopters: age, past activity of early adopters

Preference similarity: mean similarity between early adopters
[Sharma and Cosley, 2015]



100
90
80
70
60
50
Percent accuracy
all
temporal
all-temporal
reshare
structural
similarity
root
Goodreads
Flickr
Last.fm
Twitter
Dataset

A single temporal feature may be sufficient

time_5—time taken for an item to receive 5 adoptions

Even a model with this single temporal feature

- achieves more than 70% accuracy on all datasets

- accounts for nearly 97% of the accuracy of the full model for each dataset

Other features have inconsistent effects on popularity

If you fit a single-feature logistic regression:

For **12 of the 25 features**, the coefficient term flips between being positive and negative across models fit on different datasets.

E.g. Higher density—number of edges in the subgraph of early adopters

--higher popularity on Flickr (β coefficient=0.04),

-- lower popularity on Last.fm (β coefficient=-0.09).

What did we learn?

Temporal features **dominate** other features.

Temporal features also **generalize** across diverse item domains, unlike other features.

Early popularity leads to final popularity.

If an item quickly gains early adoptions, then it is more likely to be popular eventually.

But why?

How do they become popular?
What properties of items, early adopters and their social networks matter?

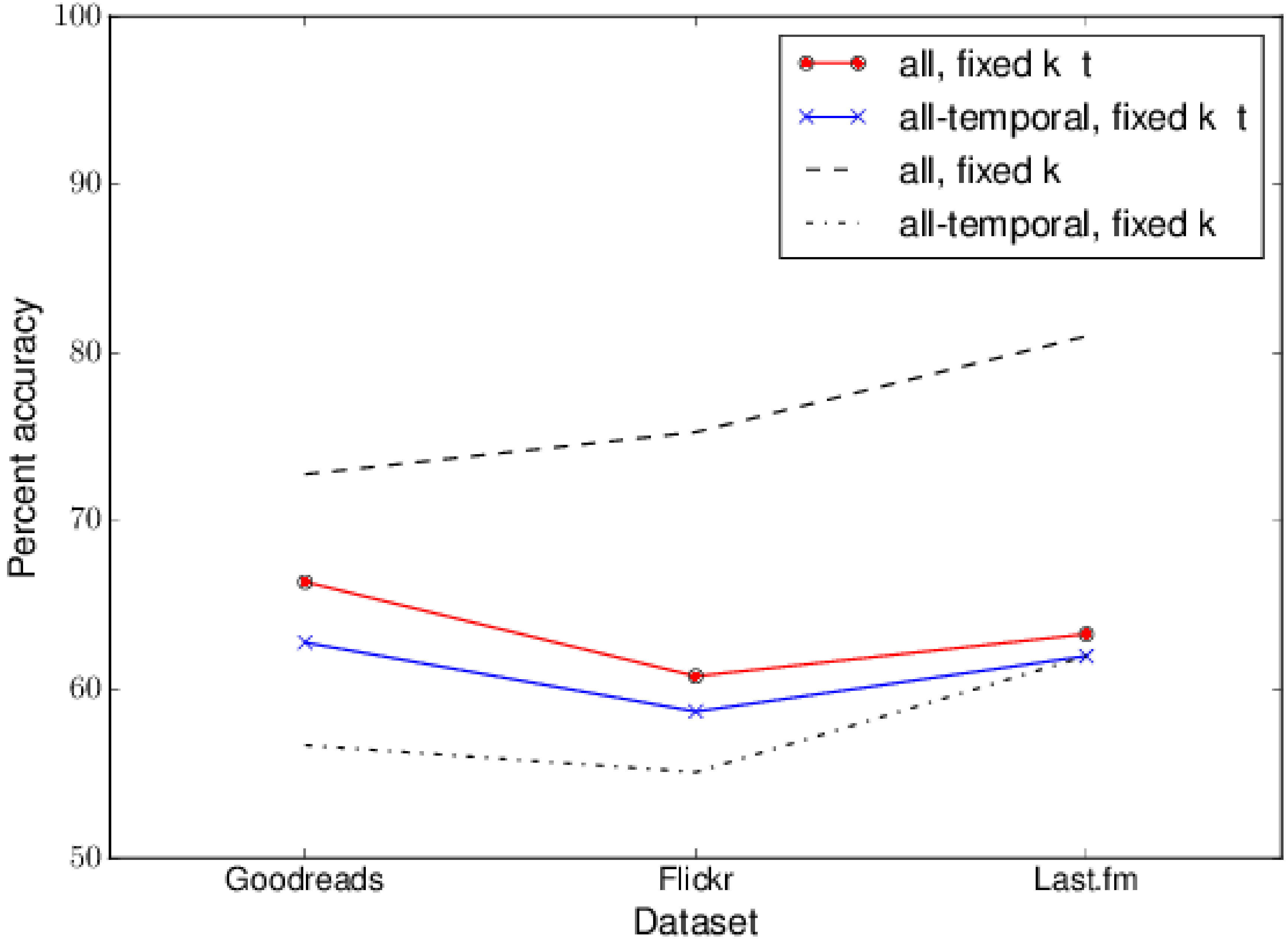
Temporally-matched Balanced Classification

Control for average rate of early adoption.

Problem:

Among items with exactly *k* adoptions at the end of a fixed time period t, which ones would be higher than the median popularity at a later time T?

Expect temporal features to play a lesser role.


100
90
80
70
60
50
Percent accuracy
Goodreads
Flickr
Last.fm
Dataset
all, fixed k t
all-temporal, fixed k t
all, fixed k
all-temporal, fixed k

Conclusion

Prediction with peeking ≠ Understanding how items gain popularity

Temporal features can be used for prediction, they even generalize across item domains.

But prediction tells us little about how items become popular.

Temporally matched formulation can be a way forward to study the effect of early adopter features.

Prediction with peeking ≠ Understanding how items gain popularity

Temporally matched formulation can be a way forward to study the effect of early adopter features.

thank you!

Amit Sharma (@amt_shrma)
www.amitsharma.in