Winter in the Twitterverse:

Addressing the Cold Start Problem

Gregory Marx

Problem

How do we recommend content to new users?

Problem

How do we recommend content to new users?

This is the so-called "cold start" problem.

What we know

@Buddy the Elf

What we know

@Buddy the Elf

Age: 50

What we know

@Buddy the Elf

Age: 50

Gender: Male

What we know

@Buddy the Elf

Age: 50

Gender: Male

Likes: Christmas, Santa, Lights, Trees, Presents, Candy, Singing, Syrup

The data

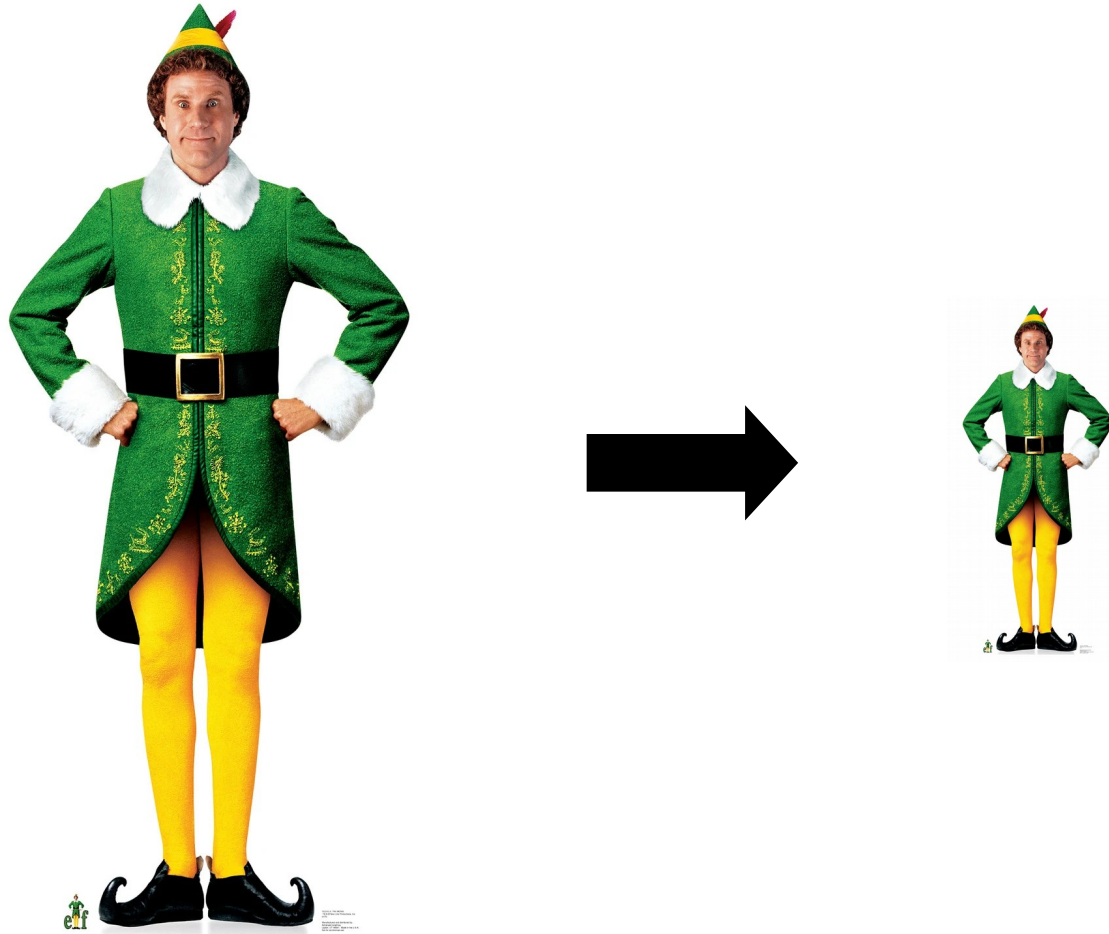- 2.3 million users

- 6,000 potential recommendations

The data

- 2.3 million users

- 6,000 potential recommendations

- 73 million data points

- 100,000 features per data point

The data

- 2.3 million users

- 6,000 potential recommendations

- 73 million data points
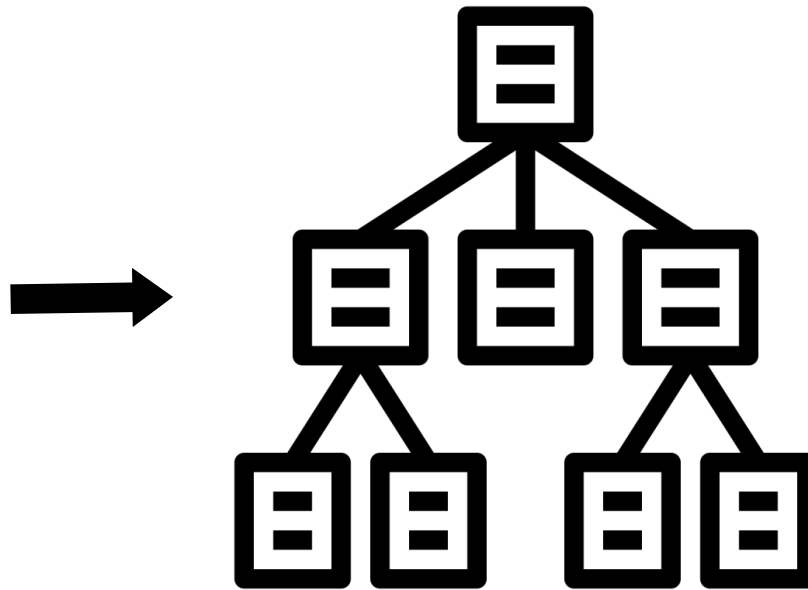
- 100,000 features per data point

➡ 27,000 GB

Methodology

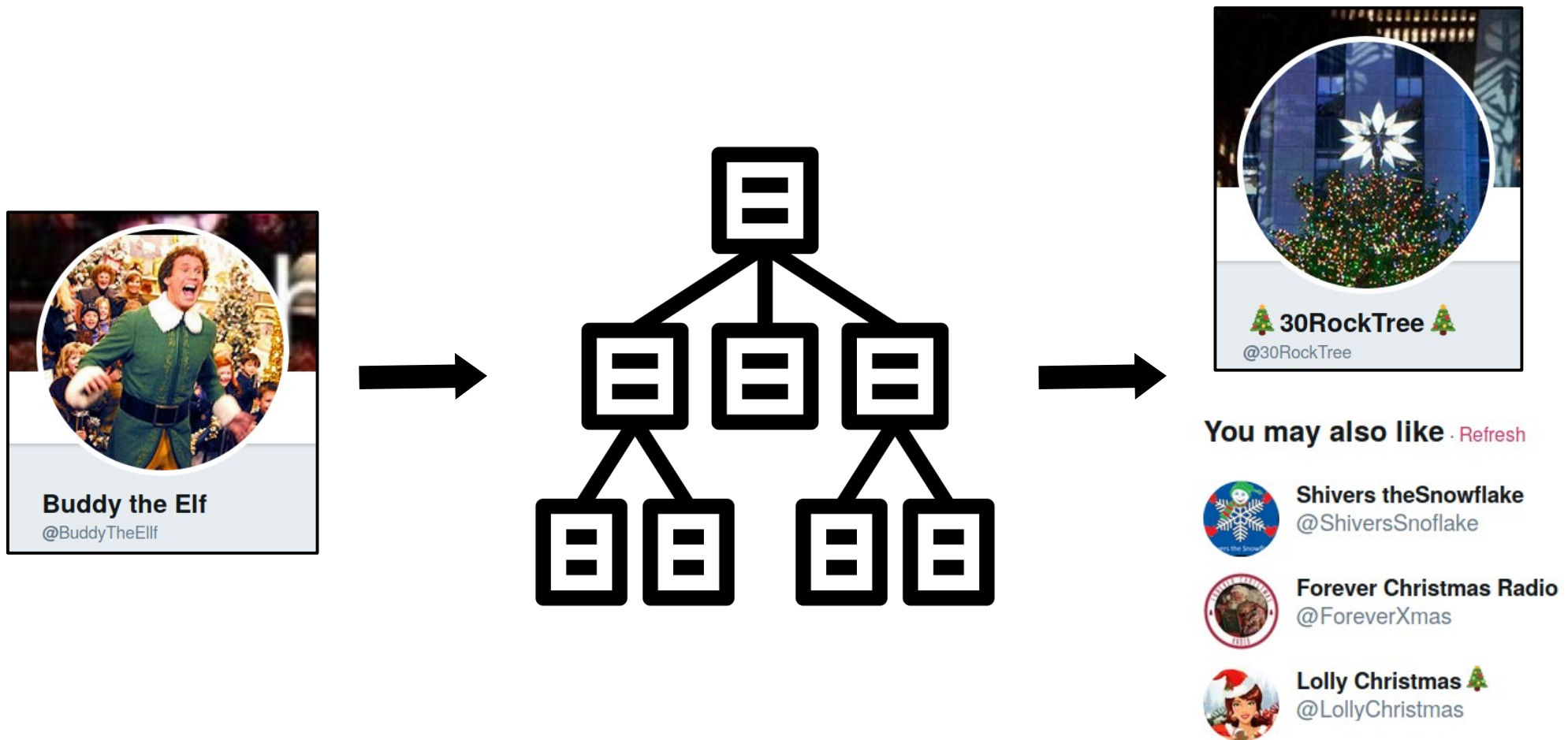# Use singular value decomposition methods to perform dimension reduction

Methodology

Implement a random forest classifier to predict likelihood of a click

Methodology

Rank items and choose top results to present to the user

Performance

Method	Mean Avg. Precision

Performance

Method	Mean Avg. Precision
Least popular	.12

Performance

Method	Mean Avg. Precision
Least popular	.12
Random	.20

Performance

Method	Mean Avg. Precision
Least popular	.12
Random	.20
Most popular	.29

Performance

Method	Mean Avg. Precision
Least popular	.12
Random	.20
Most popular	.29
My model	.35

What's next?

Buddy loved our recommendation!

What's next?

Buddy loved our recommendation!

How can we follow up with more content?

What's next?

@30RockTree

What's next?

@30RockTree

@nycballet

@FAOSchwarz

Business applications

- **Recommendation systems are used everyday by top tech companies like Netflix and Amazon**

Business applications

- Recommendation systems are used everyday by top tech companies like Netflix and Amazon

- Classification models have applications in countless industries including finance, healthcare, and manufacturing

Thank you!

Gregory Marx, PhD

gregorymarx85@gmail.com

in marxg

marxg