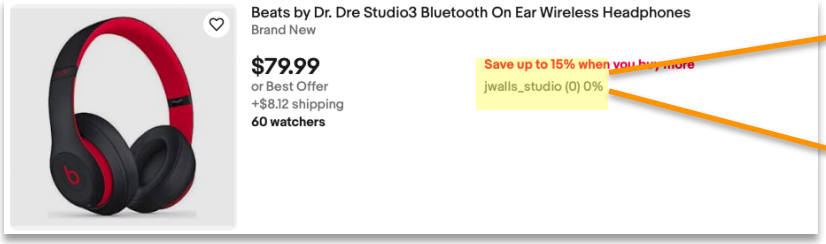
Does Online Anonymous Reputation Matter?

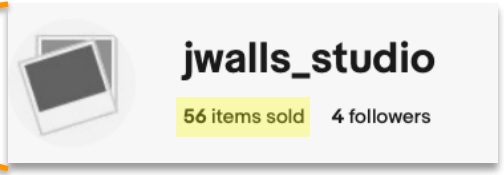
Alejandro Cuevas – acuevasv@andrew.cmu.edu

Nicolas Christin – nicolasc@andrew.cmu.edu

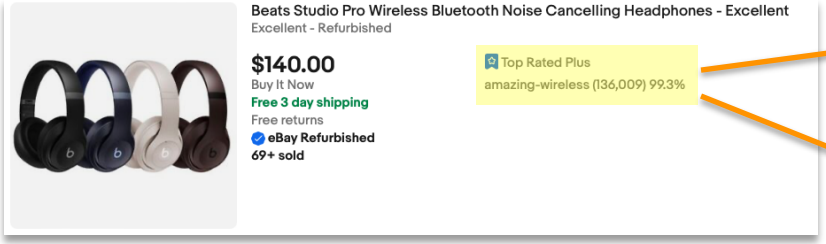
CyLab Carnegie Mellon University
Security and Privacy Institute

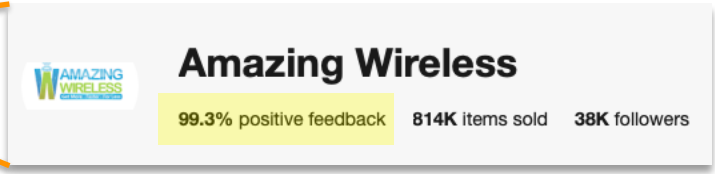
Suppose you want to buy new headphones...





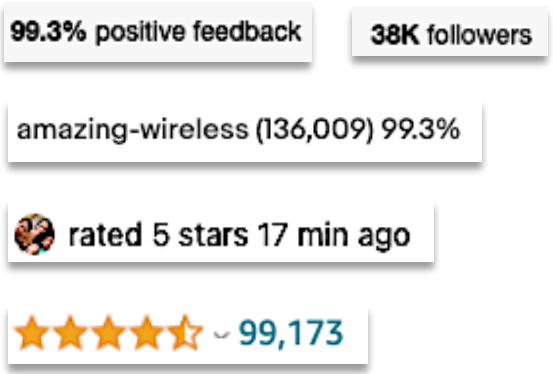
No feedback

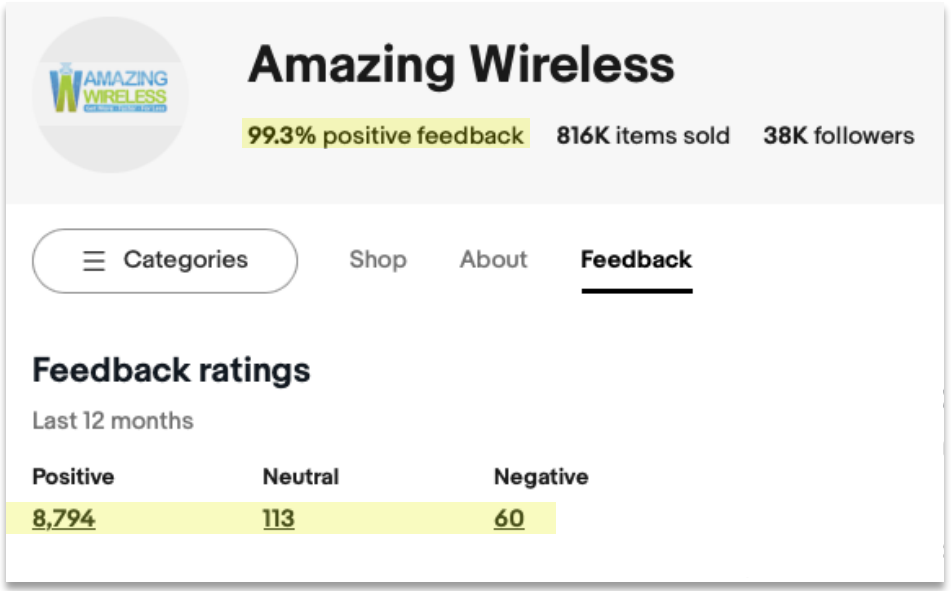




99.3% positive feedback

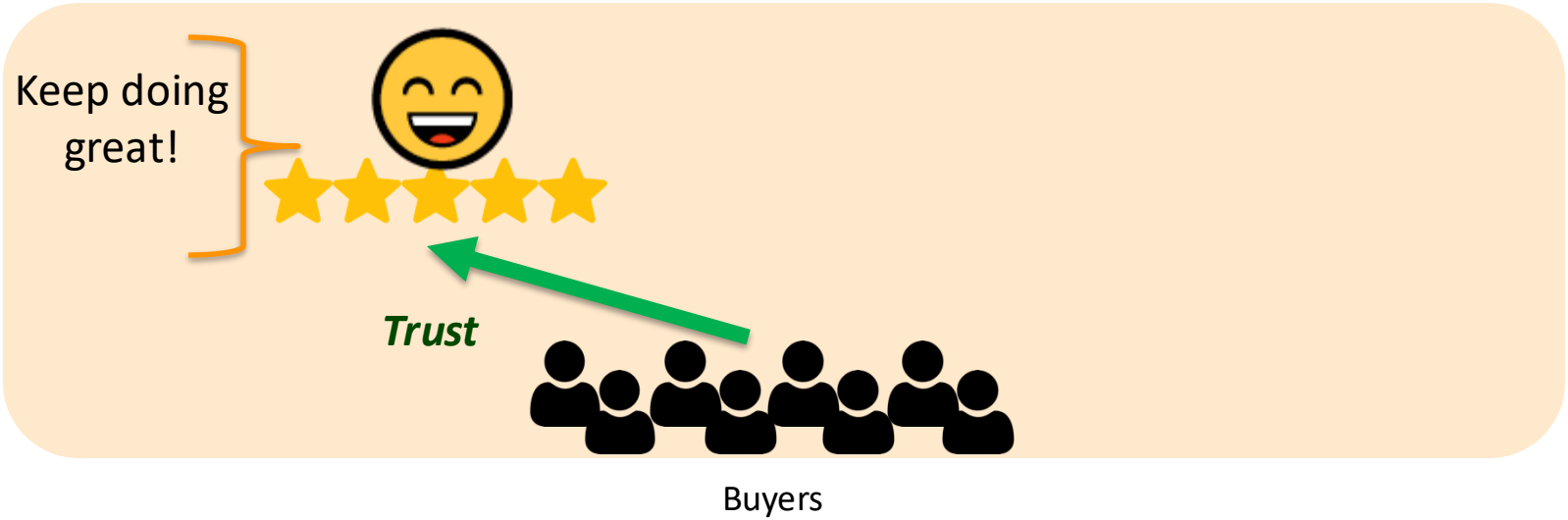
Reputation signals facilitate trust between buyers and sellers





Better reputation should lead to ***more*** trust and ***more*** sales

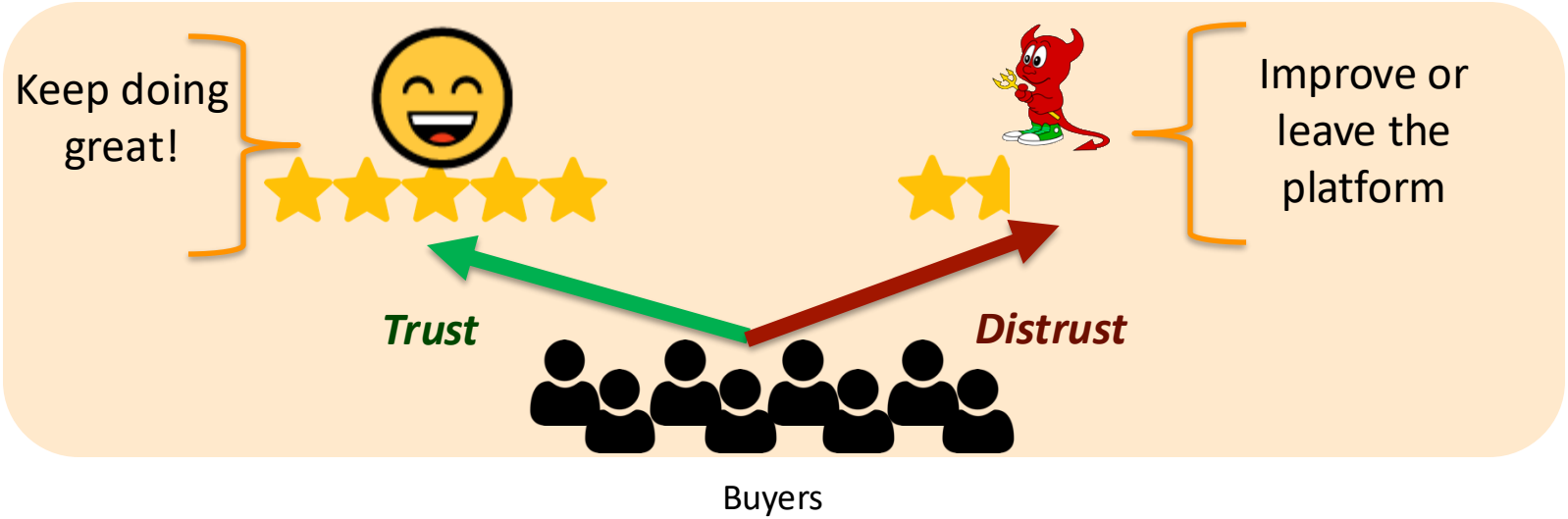
★★★★★ **Great Headphones at a great price!**



Bad reputation should lead to ***less*** trust and ***less*** sales

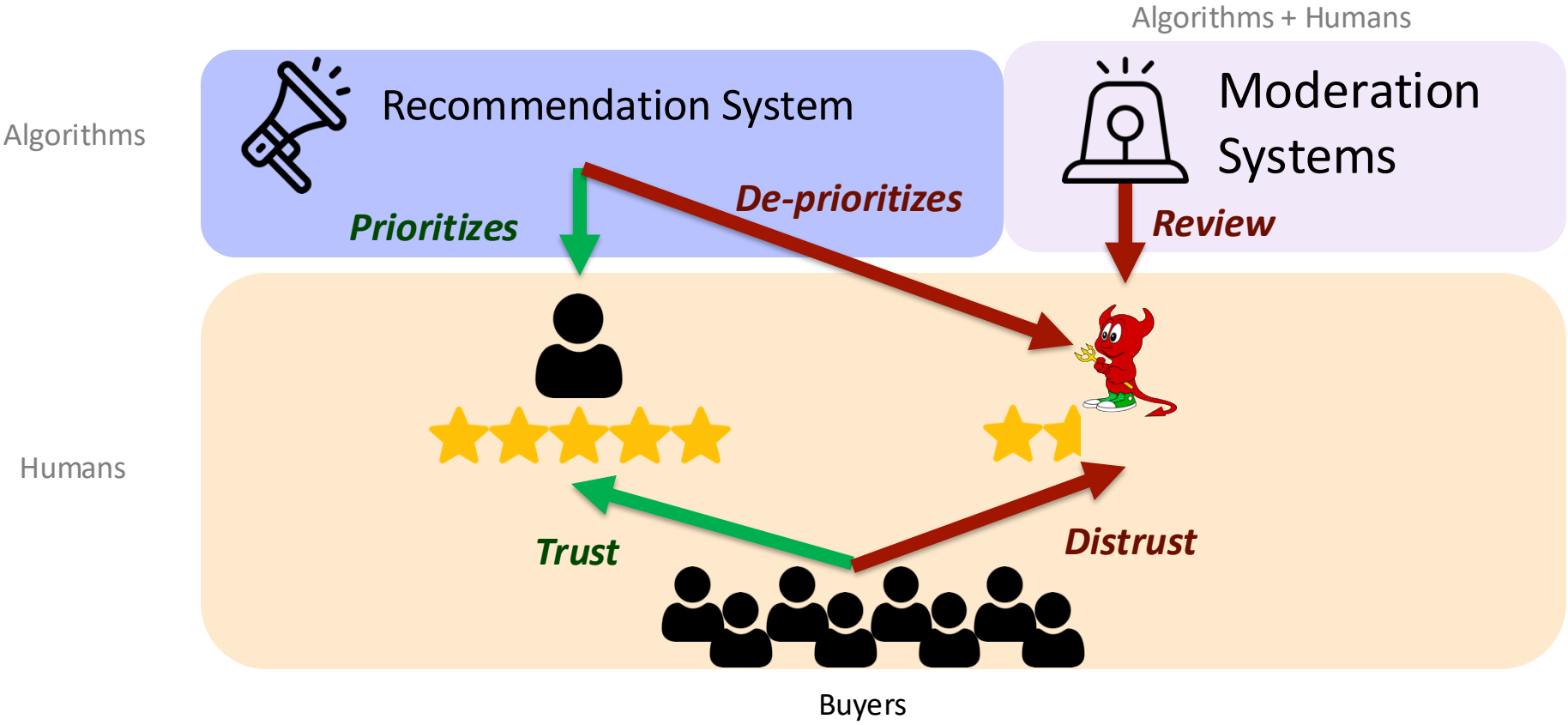
★★★★★ **Great Headphones at a great price!**

★☆☆☆☆ **DO NOT BUY THESE!!!! They are trying to trick you!**

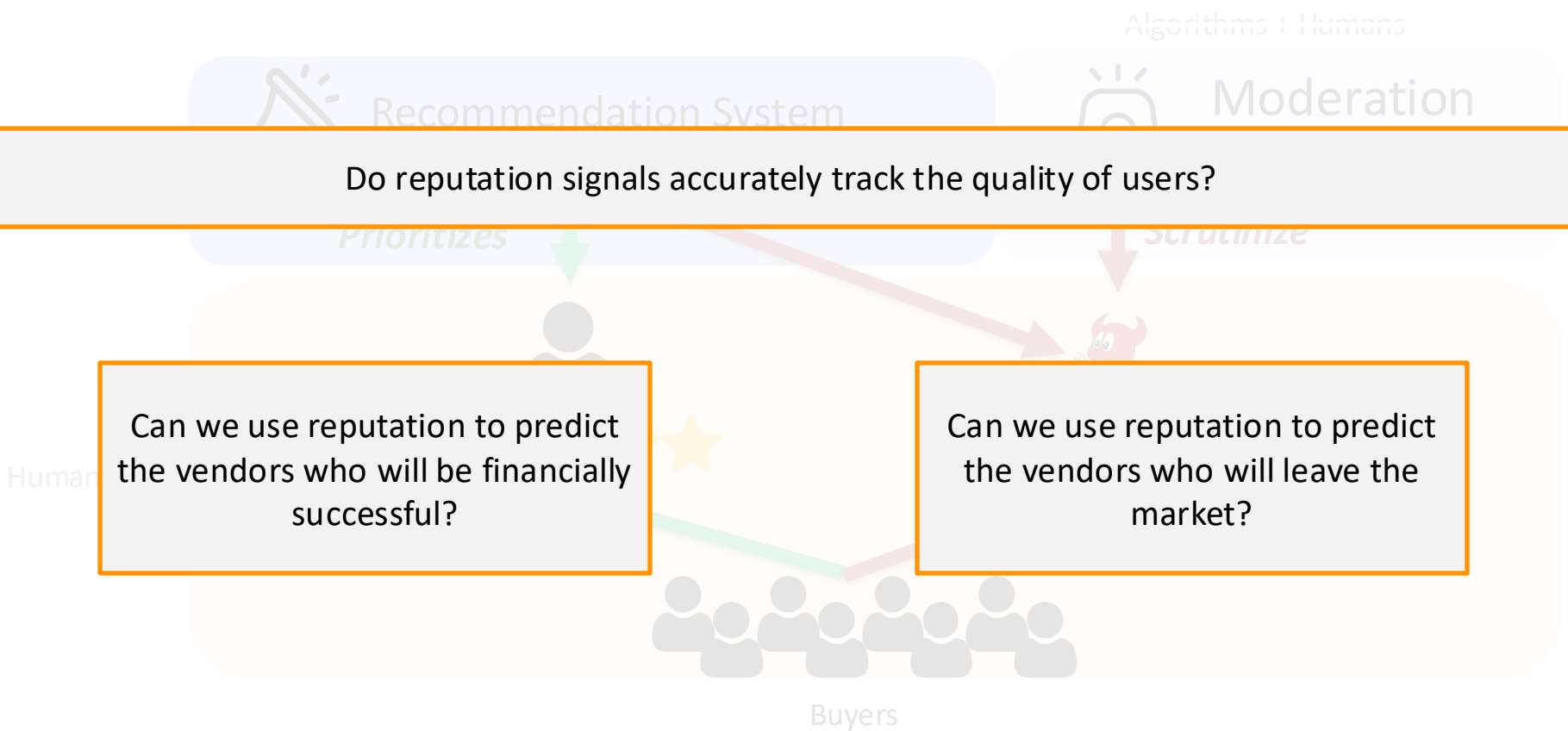


Reputation signals are everywhere

Reputation signals feed into other systems



Reputation signals feed into other systems



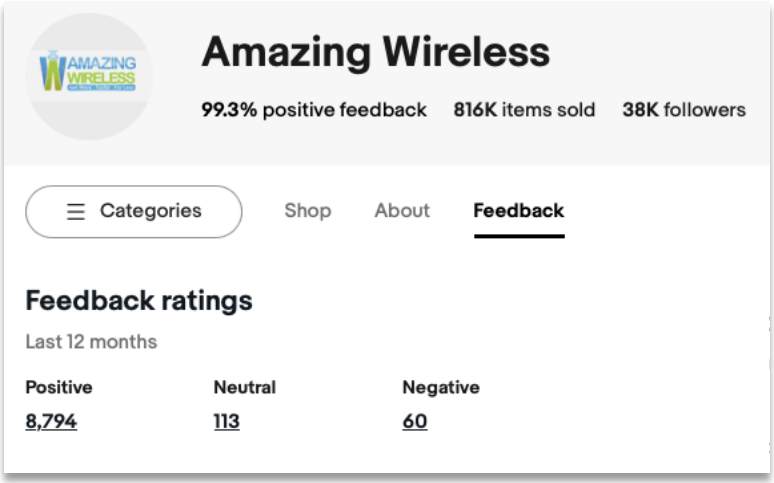
Do reputation signals accurately track the quality of users?

Can we use reputation to predict the vendors who will be financially successful?

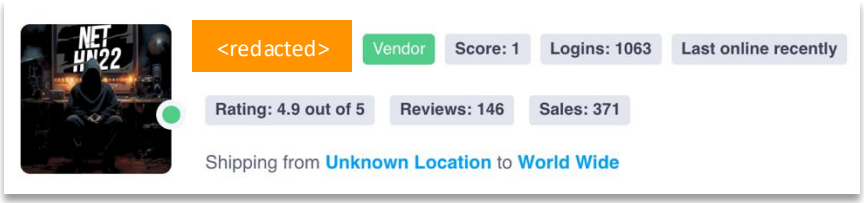
Can we use reputation to predict the vendors who will leave the market?

Online anonymous markets (OAMs) behave just like the traditional markets we are used to

A vendor on eBay

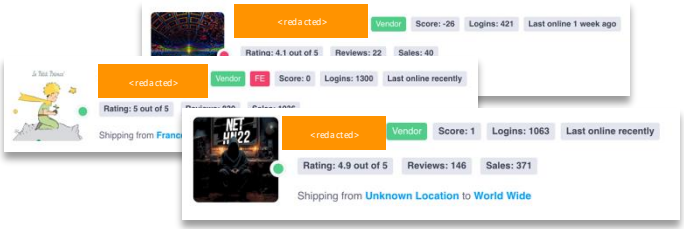


A vendor on Nemesis Market



OAMs are a great place to study reputation

1. We have validated methods to estimate sales [Cuevas et al. USENIX'22]
   - And we have a comprehensive dataset

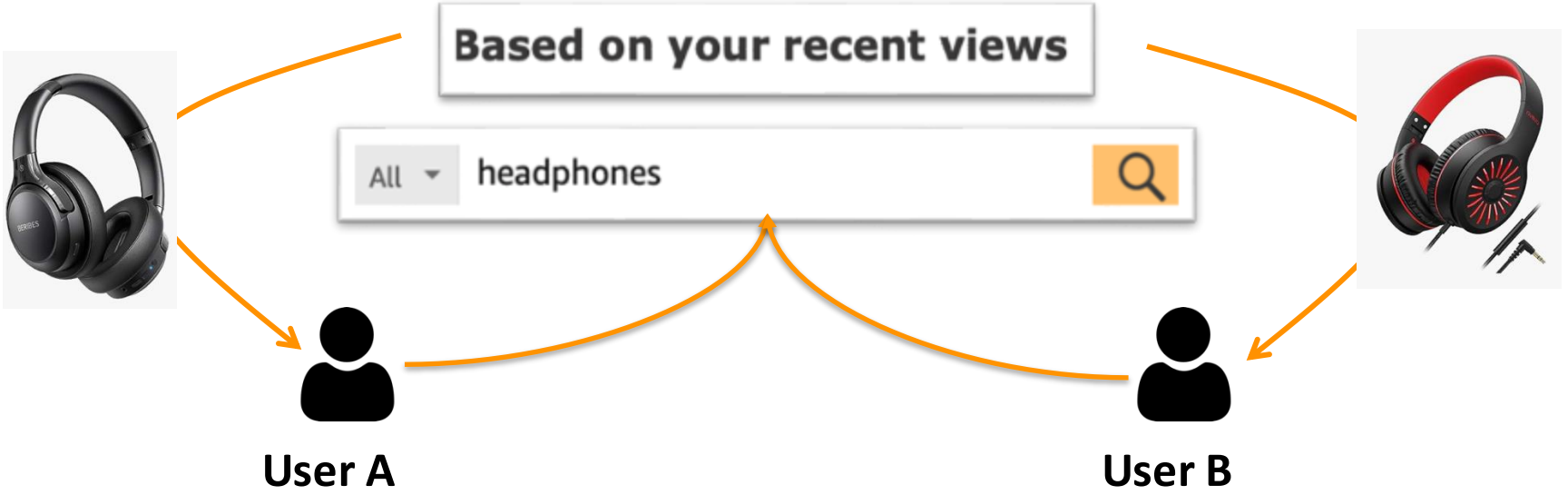


8 individual markets *(2011-2023)*
>16,000 vendors

OAMs are a great place to study reputation

1. We have validated methods to estimate sales [Cuevas et al. USENIX'22]

2. Little (if any) algorithmic interference in users' search

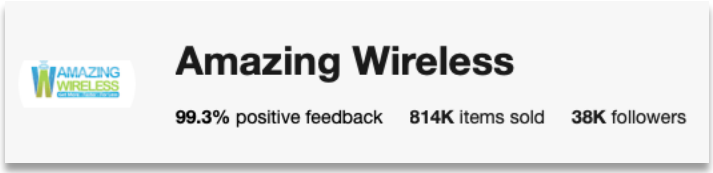


OAMs are a great place to study reputation

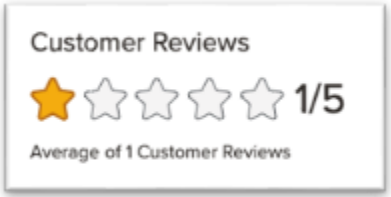
1. We have validated methods to estimate sales [Cuevas et al. USENIX'22]

2. Little (if any) algorithmic interference in users' search

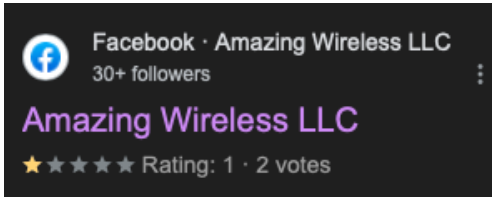
3. Few external signals about vendors



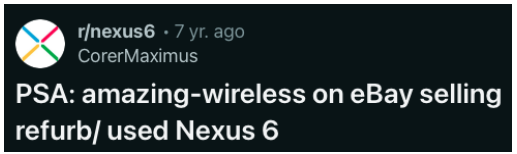
eBay



BBB



Facebook



Reddit

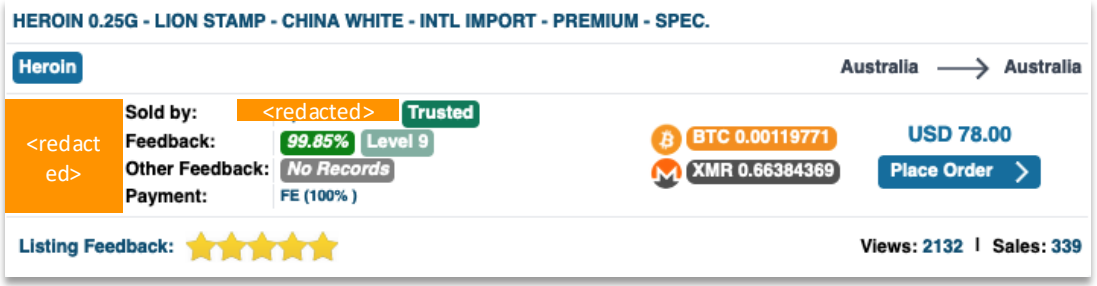
OAMs are a great place to study reputation

1. We have validated methods to estimate sales [Cuevas et al. USENIX'22]

2. Little (if any) algorithmic interference in users' search

3. Few external signals about vendors

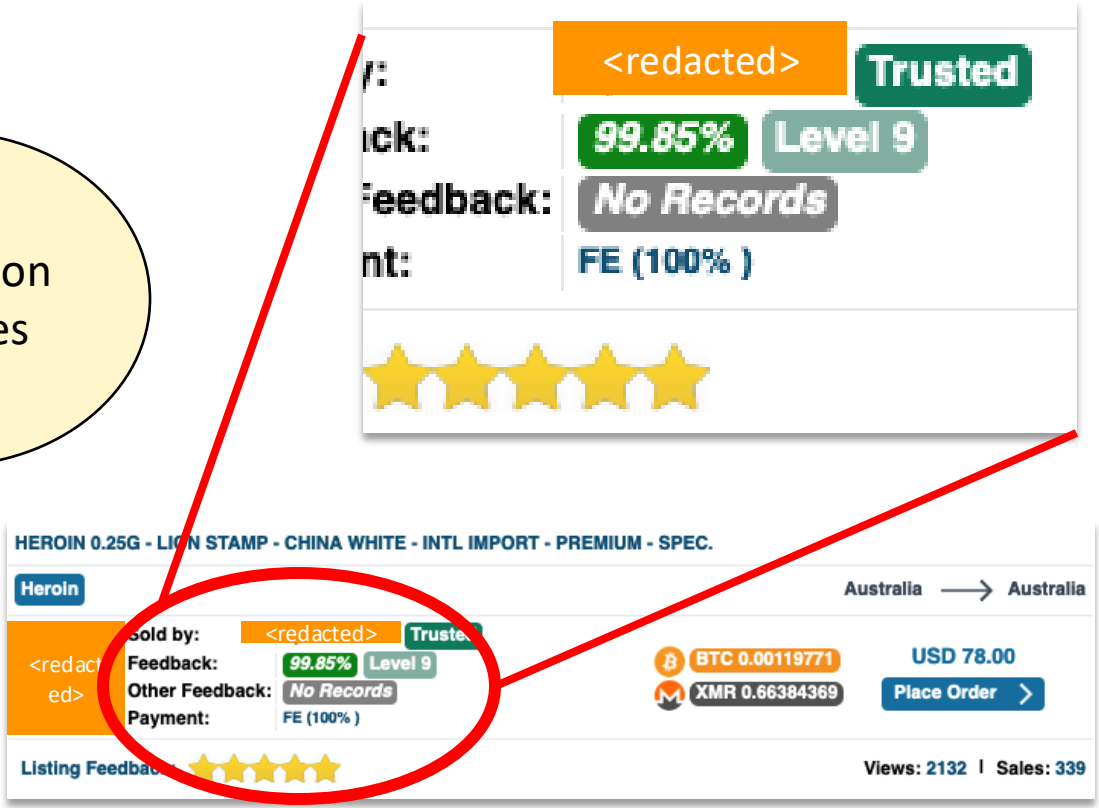
4. Reputation is **EXTREMELY** important and users often leave reviews



Reputation features include any features derived from ratings by other users

Reputation Features

E.g.,

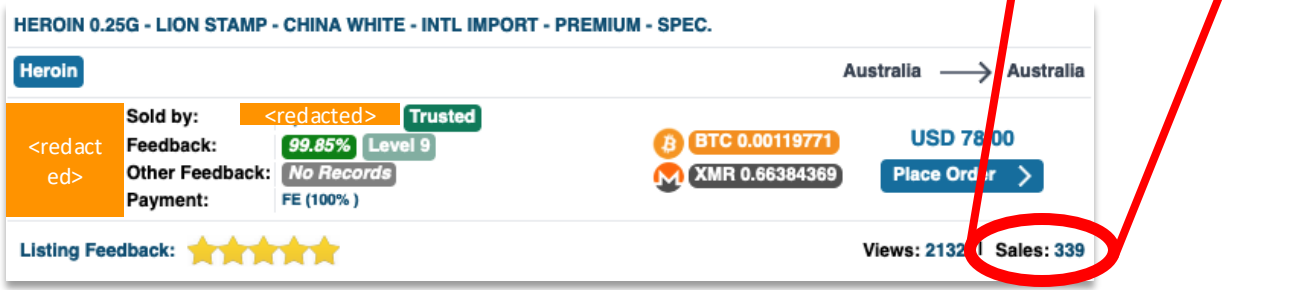


Revenue features include data about sales, such as the number of sales or rolling averages

Reputation Features

Revenue Features

E.g.,



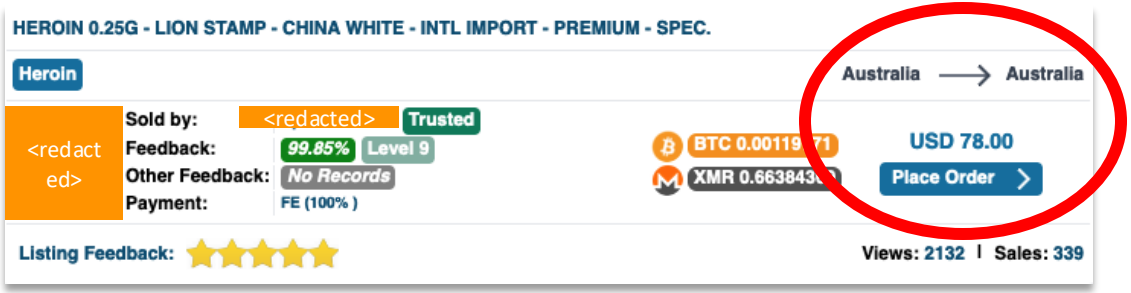
Listing features include things like the category of goods, the price, or where it ships from

Reputation Features

Revenue Features

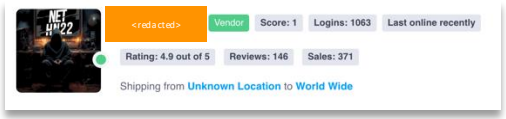
Listing Features

E.g.,



Does reputation predict the financial success?

1



Grab all available vendors

2

Train model

Month 1

...

Month *n*

W_1	W_2	W_3	W_4
W	W	W	W
W_{n-3}	W_{n-2}	W_{n-1}	W_n

Weeks

3

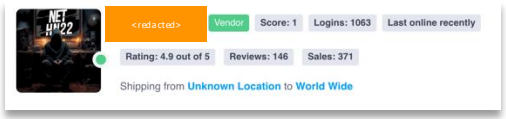
Predict earning bracket

Wealth Quartiles

Q4	*Top 25%*
Q3	
Q2	
Q1	*Bottom 25%*

Does reputation predict the financial success?

1



Grab all available vendors

2

Train model

Month 1

...

Month *n*

W_1 W_2 W_3 W_4

W W W W

W_{n-3} W_{n-2} W_{n-1} W_n

Weeks

3

Predict earning bracket

Wealth Quartiles

Q4 — *Top 25%*

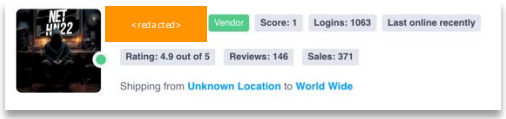
Q3

Q2

Q1 — *Bottom 25%*

Does reputation predict the financial success?

1



Grab all available vendors

2

Train model

Month 1

...

Month *n*

W_1 W_2 W_3 W_4

W W W W

W_{n-3} W_{n-2} W_{n-1} W_n

Weeks

3

Predict earning bracket

Wealth Quartiles

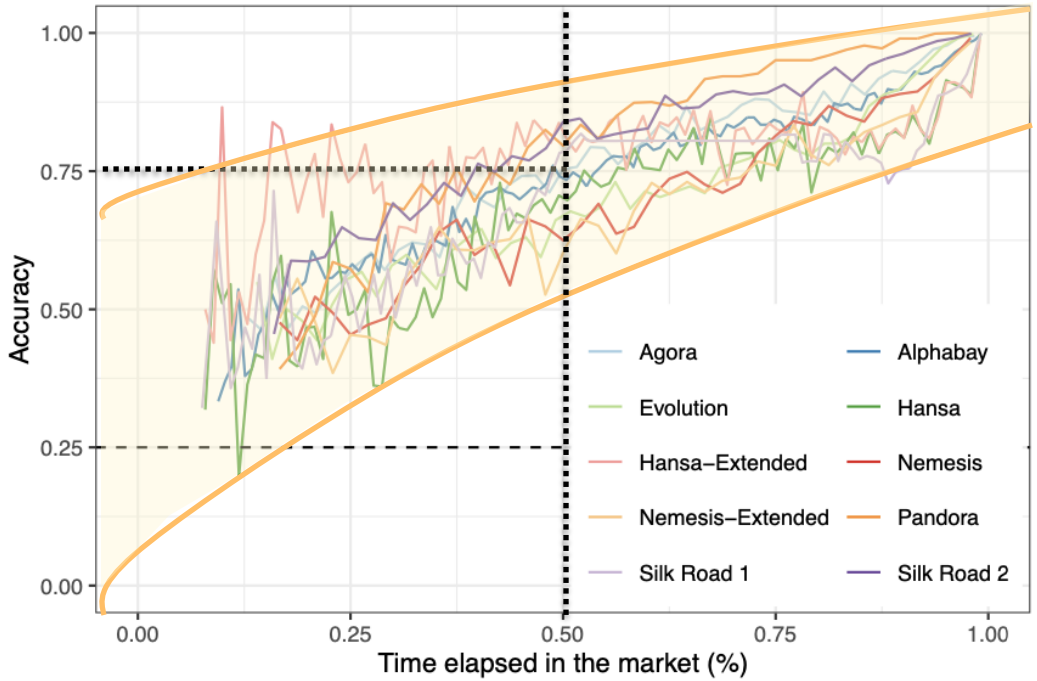
Q4 — *Top 25%*

Q3

Q2

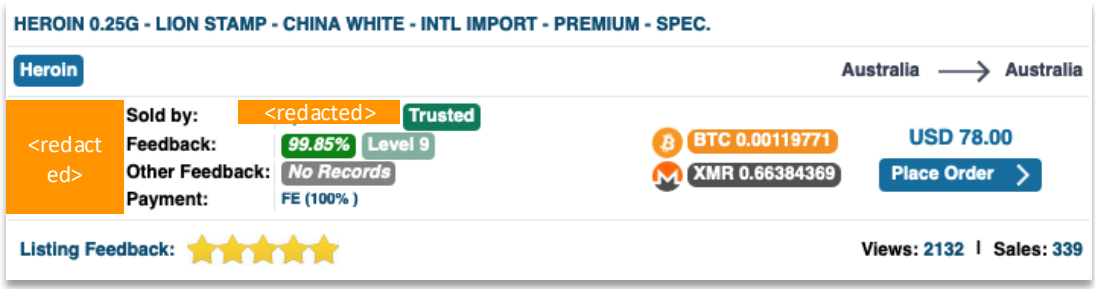
Q1 — *Bottom 25%*

We can predict success quite well using simple a model



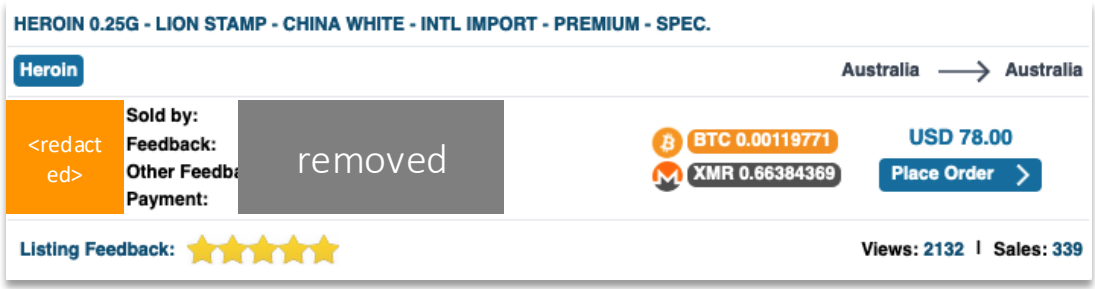
Halfway through a market's lifetime we had ~75% accuracy

If we remove reputation features...



Excluded Feature(s)		Avg. Accuracy Decrease			
Feature Set 1	Feature Set 2	Min.	Max.	Mean	Std.

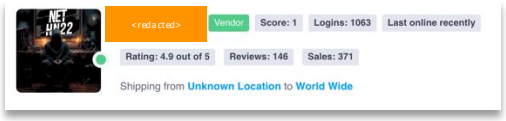
We see that reputation alone has little impact on the prediction



Excluded Feature(s)		Avg. Accuracy Decrease			
Feature Set 1	Feature Set 2	Min.	Max.	Mean	Std.
Reputation	–	<0.01	<0.01	<0.01	<0.01

Does reputation predict financial failure?

1



Grab all available vendors

2

Train model

W_1 W_2 W_3 W_4 W_5

1 month

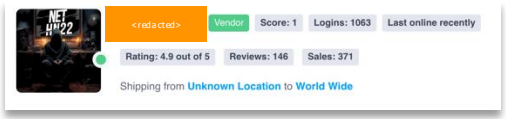
3

***Predict**: Did vendor have their last sale? (y/n)*

High Predicted Likelihood

Does reputation predict financial failure?

1



Grab all available vendors

2

Train model

3

W_1

W_2 W_3 W_4 W_5

W_6 W_7 W_8 W_9

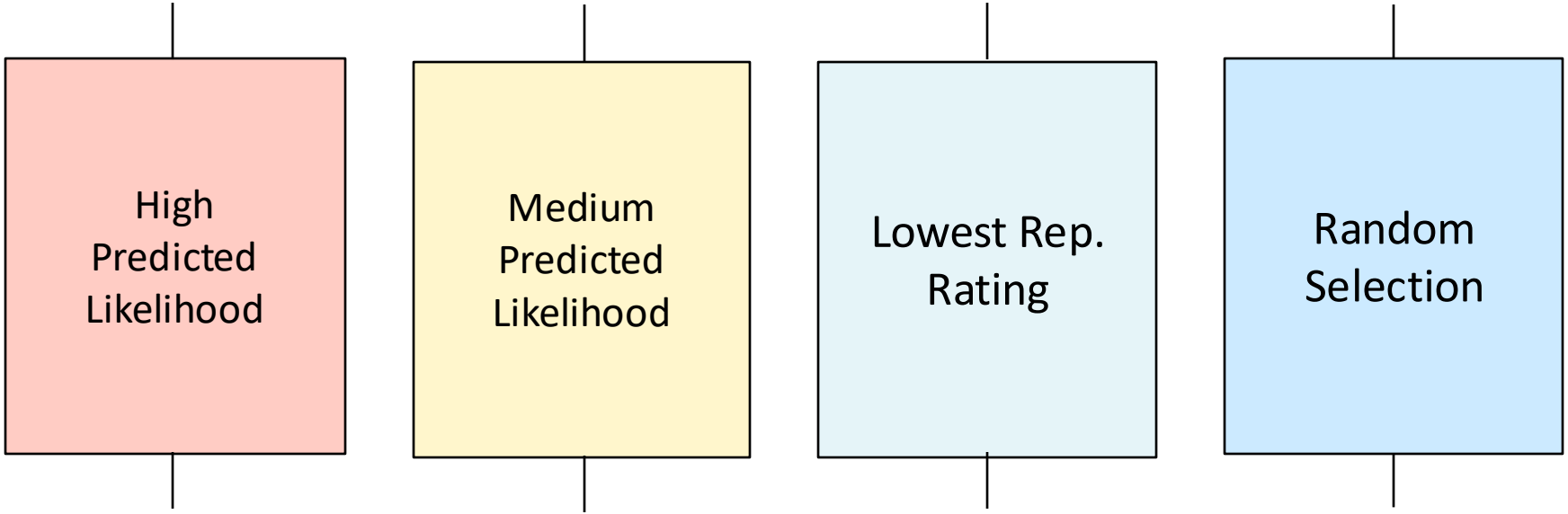
W_{10} W_{11} W_{12} W_{13}

1-3 months

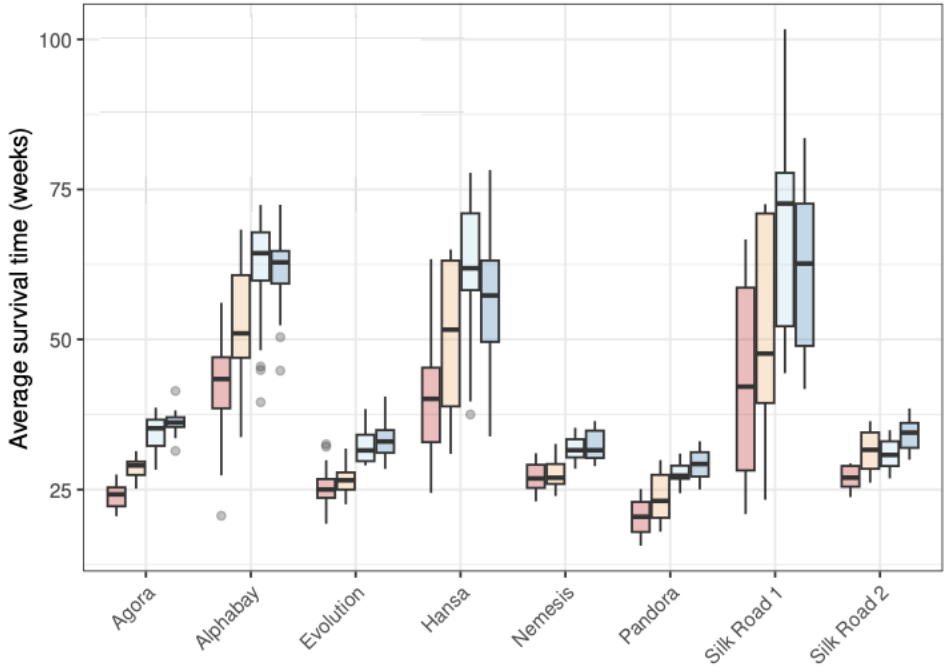
***Predict**: Did vendor have their last sale? (y/n)*

Medium Predicted Likelihood

We assemble 4 groups of *n* vendors across markets



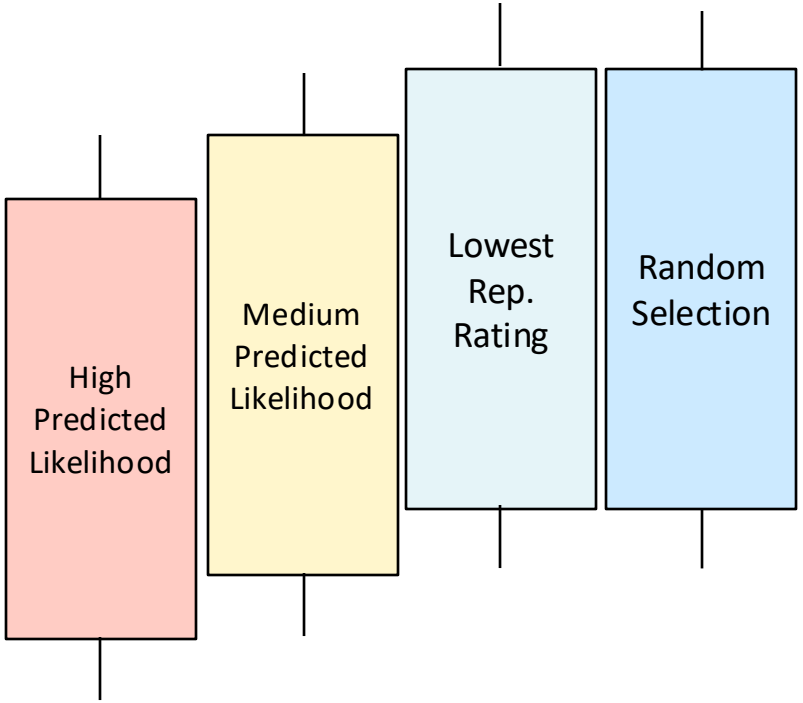
Public signals can help you identify vendors who will leave the market

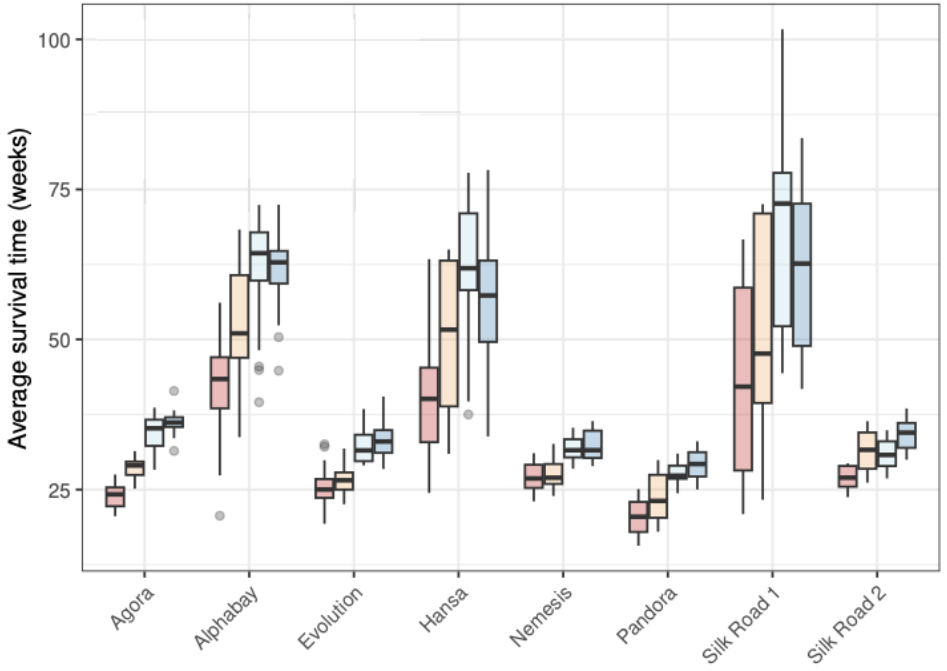


Shorter boxplot means shorter lifespan

Differing y-axis lengths are due to different market lengths

Reputation alone is not a good indicator





We can identify users who will disappear faster using a combination of signals

Reputation systems can be a powerful tools, but they have become less effective

The New York Times

Fake Reviews Are Rampant Online. Can a Crackdown End Them?

THE WALL STREET JOURNAL.

Special Off

Business U.S. Politics Economy Tech Markets & Finance Opinion Arts Lifestyle Real Estate Personal Finance Health

Black Friday Shoppers: Beware of Fake Five-Star Reviews

1. We introduce a new method to evaluate reputation systems in marketplaces

2. We demonstrate that reputation alone has little impact, but effective predictions rely on a variety of public signals

3. We highlight the need to continuously evaluate signals and suggest how we can design more effective reputation systems

Alejandro Cuevas –
acuevasv@andrew.cmu.edu
alejandrocuevas.me

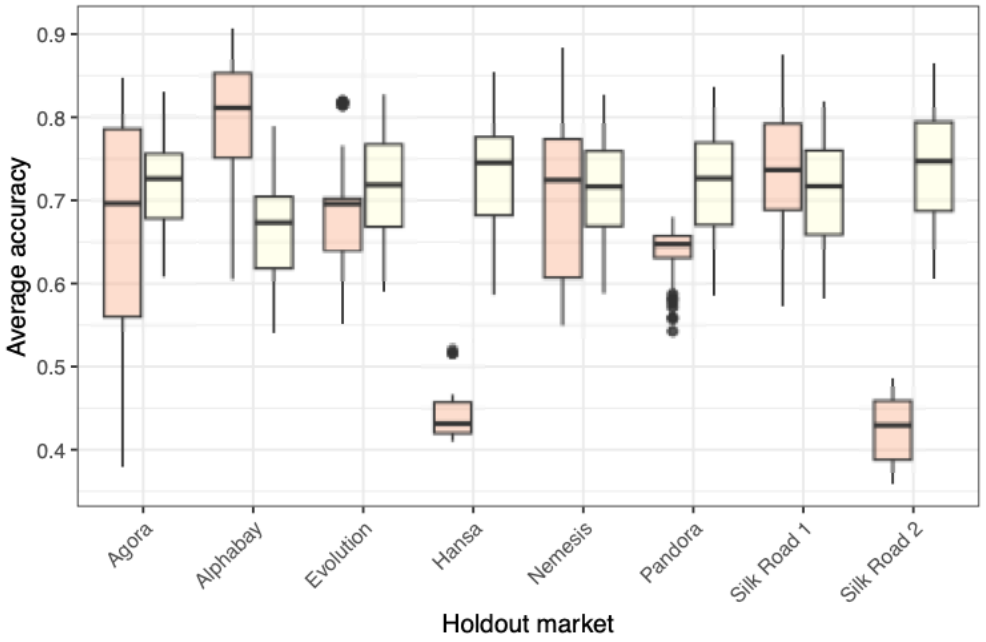
CyLab Carnegie Mellon University Security and Privacy Institute

Our model mostly generalizes across markets and time

For all markets:

Train a model on n-1 markets

Test model on 1 market (the 'holdout')

Training score

Holdout score