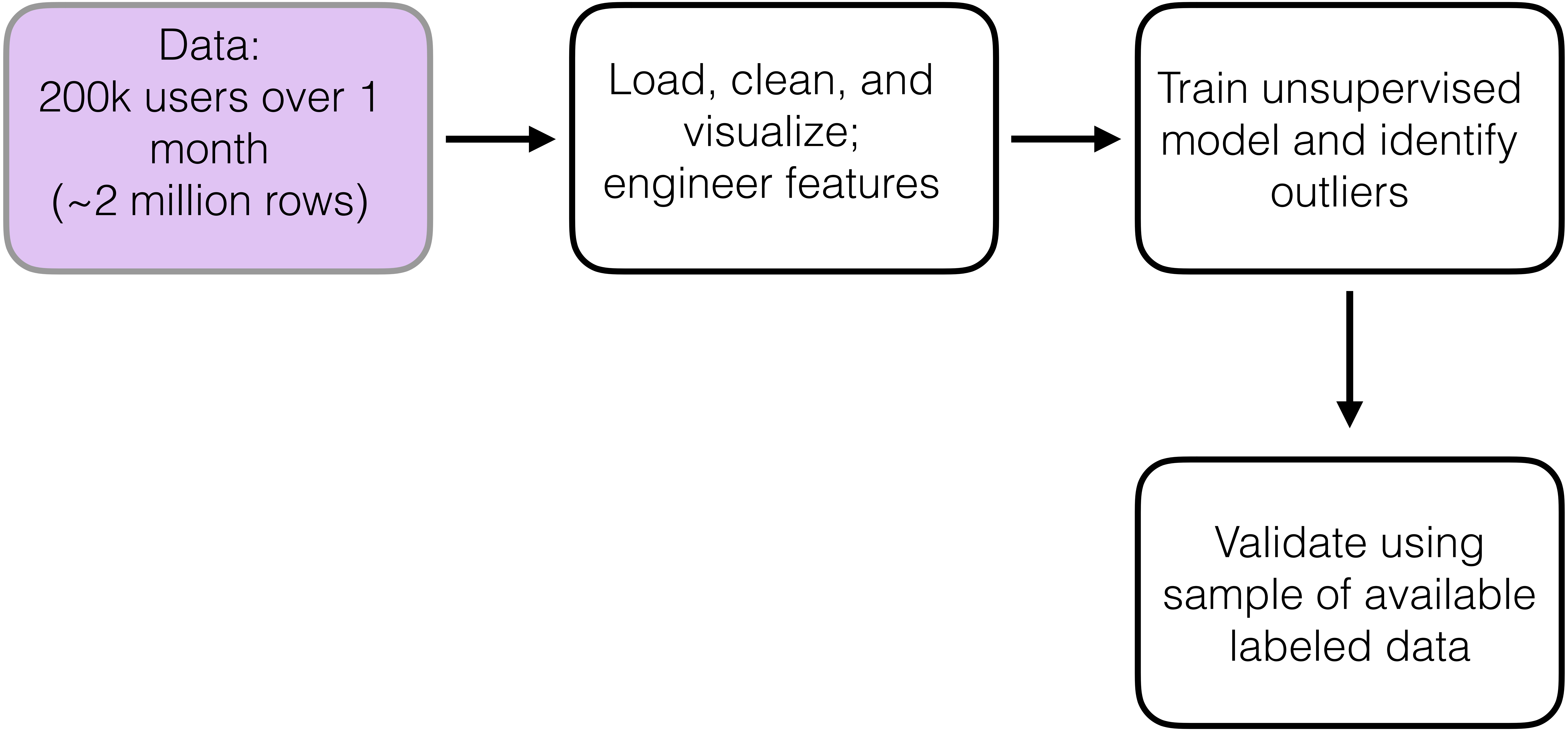
Detection of unusual travel patterns to prevent user account compromise

James Verbus

Collaboration with Castle (castle.io)

- Castle provides automated detection of compromised user accounts & hijack attempts for online businesses
- **Deliverable:** Develop a model to predict the likelihood that a new login belongs to a specific user

Data and Analysis Pipeline



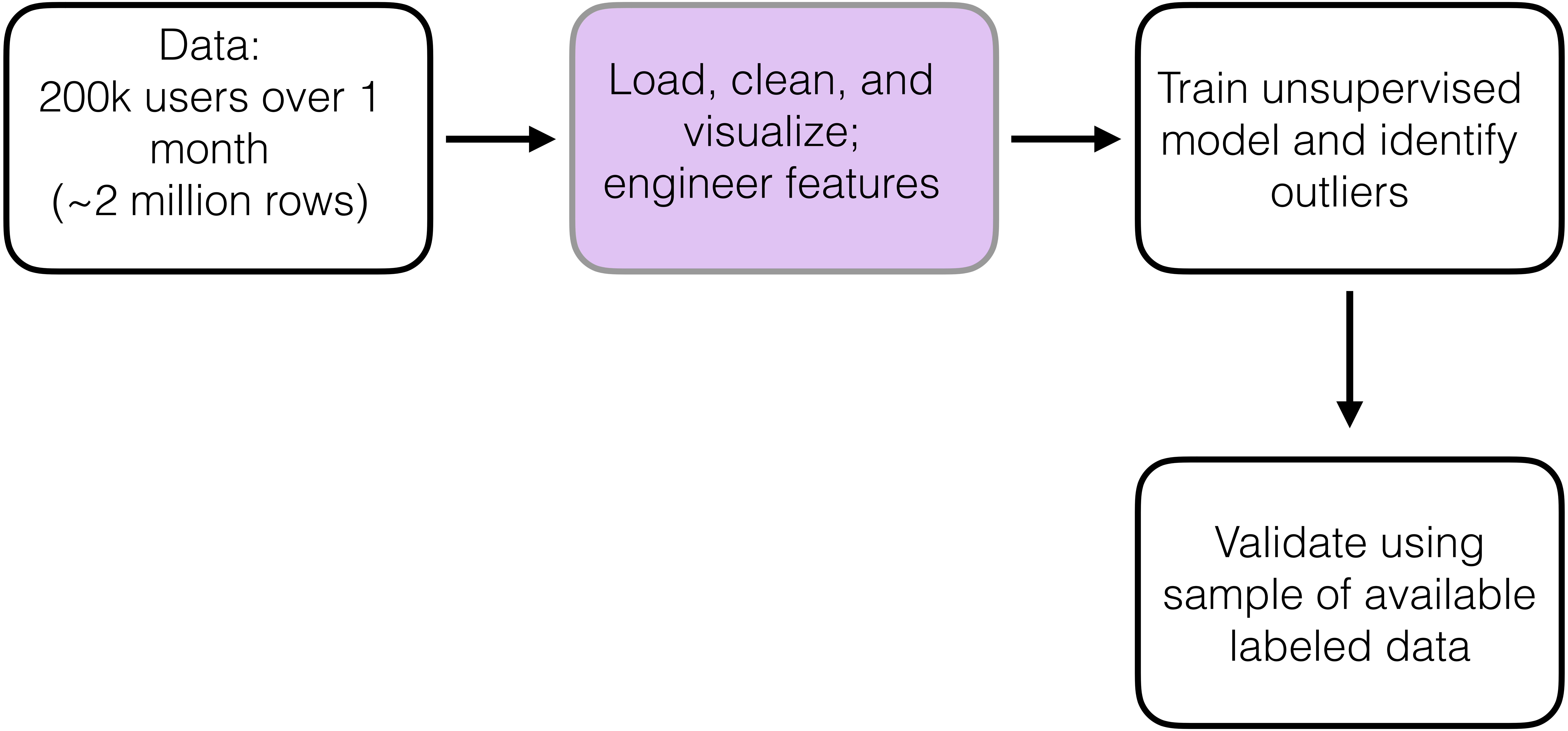
Data:
200k users over 1 month
(~2 million rows)

Load, clean, and visualize; engineer features

Train unsupervised model and identify outliers

Validate using sample of available labeled data

Data and Analysis Pipeline



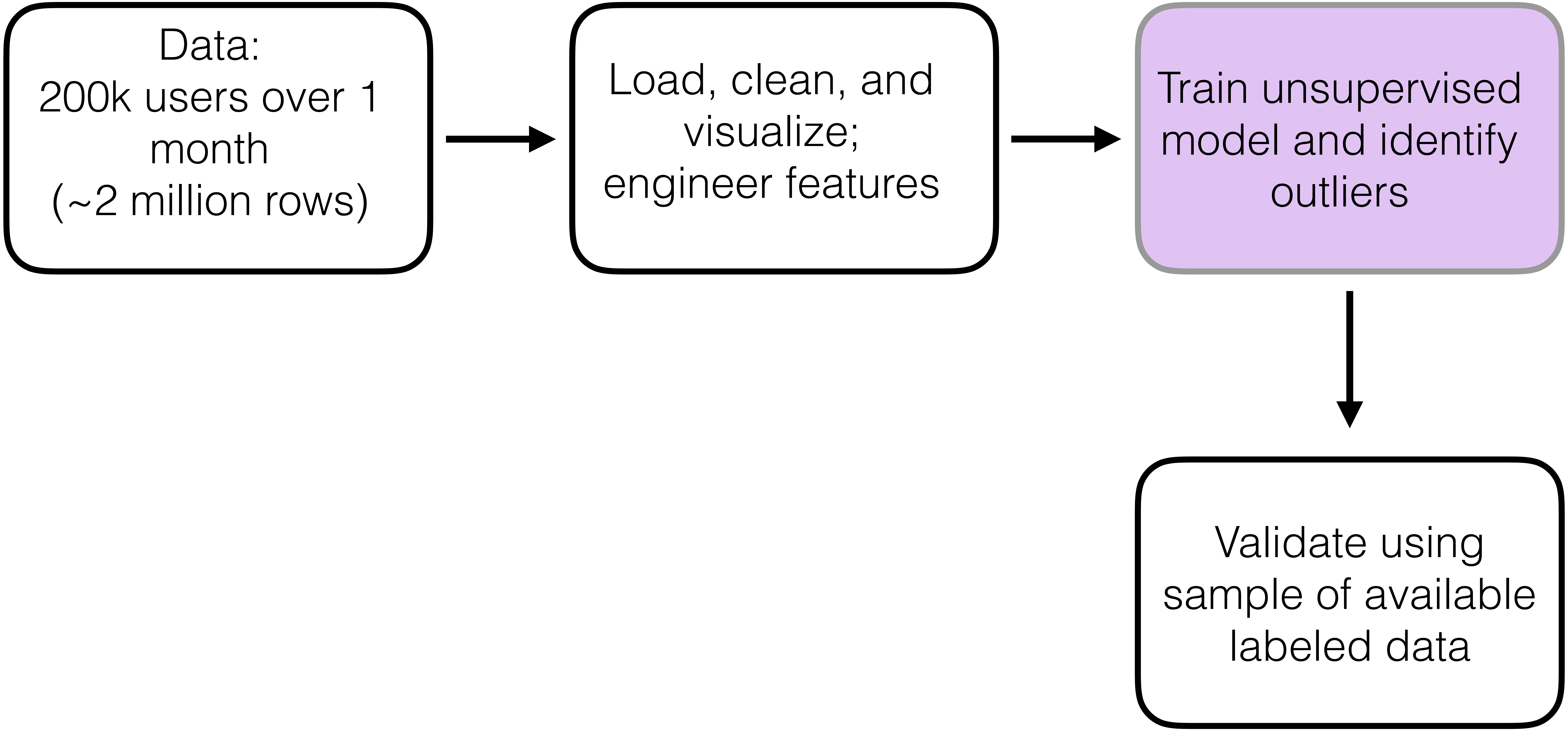
Data:
200k users over 1 month
(~2 million rows)

Load, clean, and visualize;
engineer features

Train unsupervised model and identify outliers

Validate using sample of available labeled data

Data and Analysis Pipeline



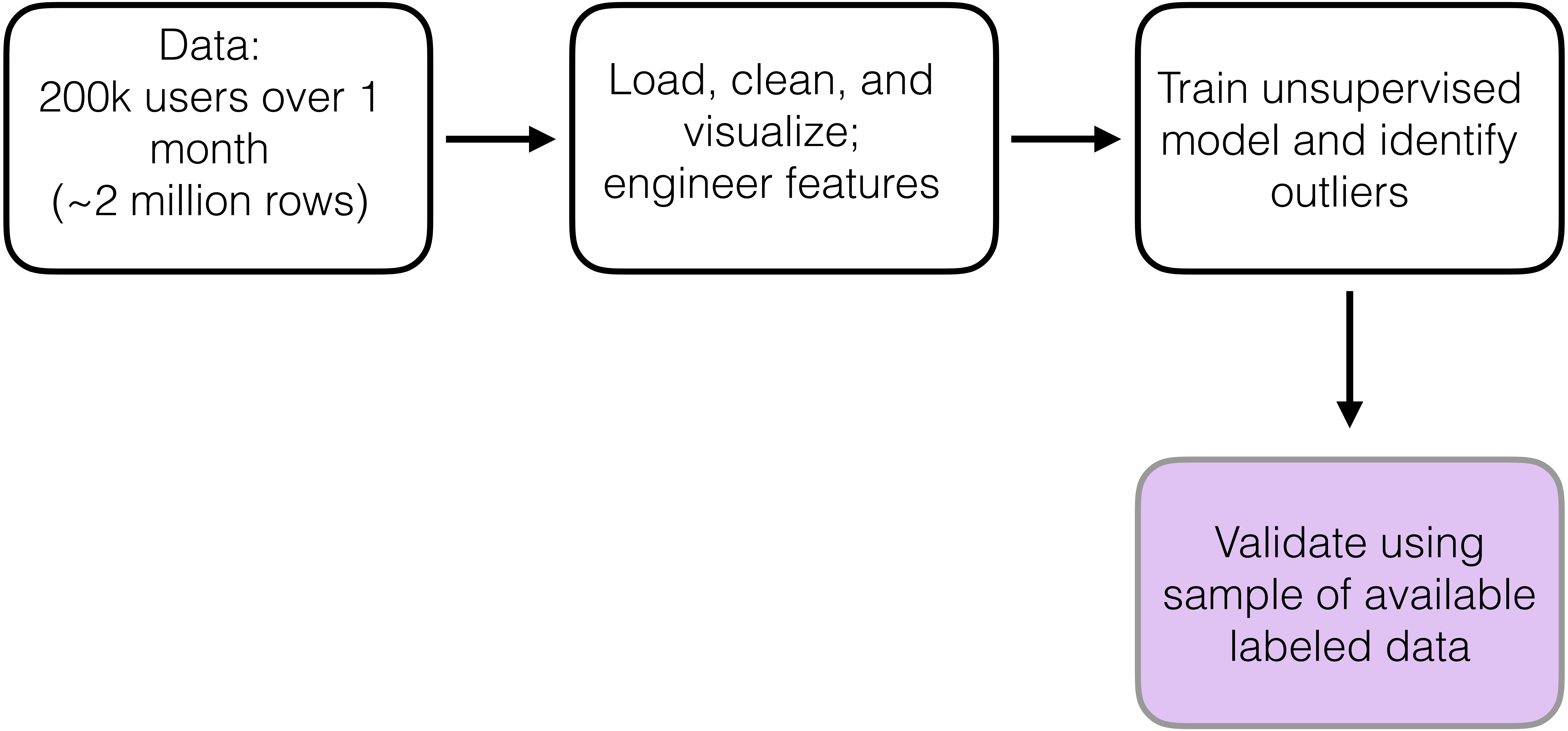
Data:
200k users over 1 month
(~2 million rows)

Load, clean, and visualize;
engineer features

Train unsupervised model and identify outliers

Validate using sample of available labeled data

Data and Analysis Pipeline



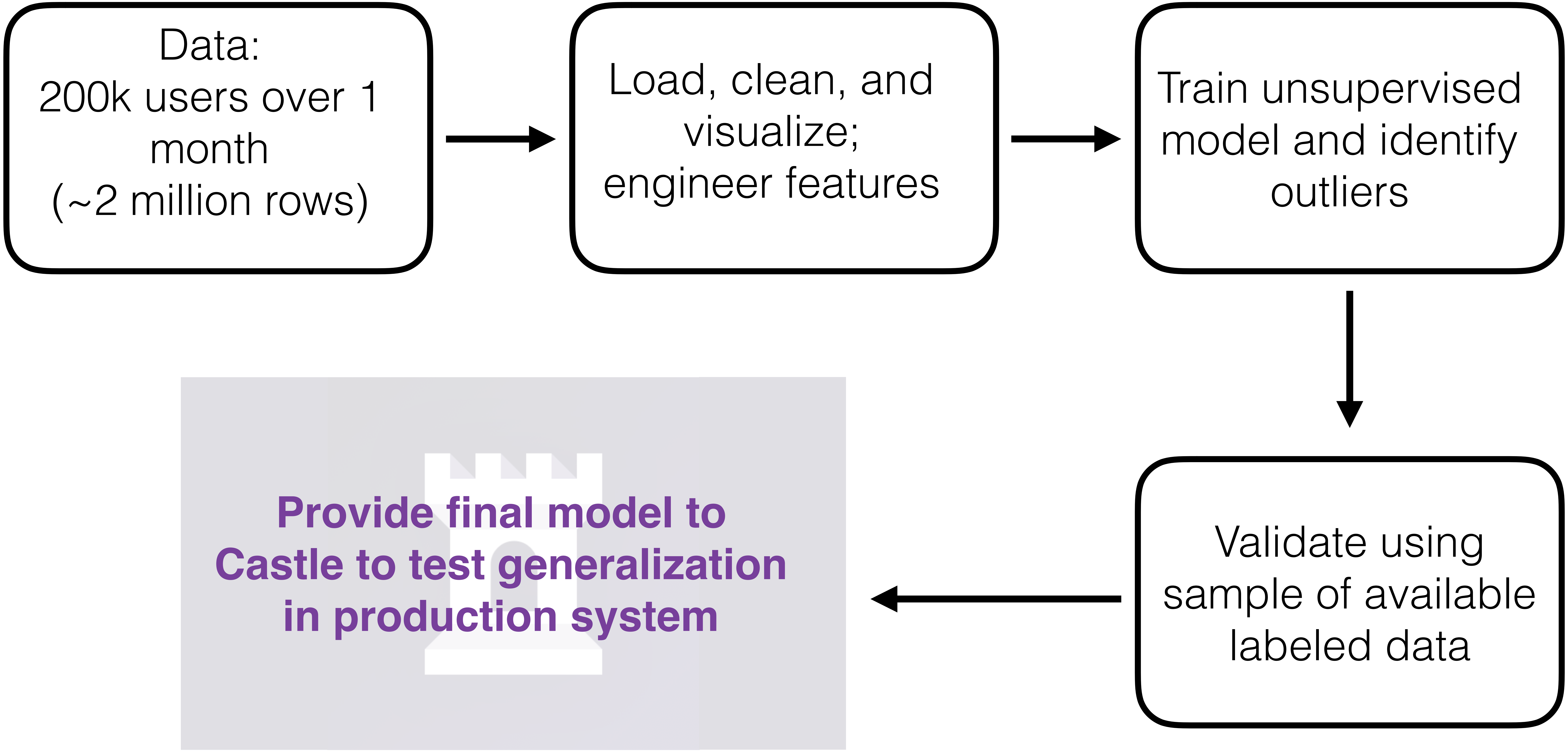
Data:
200k users over 1 month
(~2 million rows)

Load, clean, and visualize;
engineer features

Train unsupervised model and identify outliers

Validate using sample of available labeled data

Data and Analysis Pipeline



Data:
200k users over 1 month
(~2 million rows)

Load, clean, and visualize;
engineer features

Train unsupervised model and identify outliers

Validate using sample of available labeled data

Provide final model to Castle to test generalization in production system

Canada
Hudson Bay
United States
Mexico
Gulf of Mexico
Cuba
Puerto Rico
Caribbean Sea
North Atlantic Ocean
South Atlantic Ocean
Indian Ocean
Russia
Finland
Norway
Denmark
Ireland
United Kingdom
Poland
Belarus
Germany
Ukraine
France
Austria
Romania
Italy
Spain
Portugal
Greece
Turkey
Kazakhstan
Mongolia
Uzbekistan
Kyrgyzstan
Turkmenistan
China
South Korea
Sea of Japan
East China Sea
Syria
Iraq
Iran
Afghanistan
Pakistan
Nepal
India
Morocco
Algeria
Tunisia
Libya
Egypt
Western Sahara
Saudi Arabia
Oman
Yemen
Mauritania
Mali
Niger
Chad
Sudan
Burkina Faso
Guinea
Ghana
Nigeria
South Sudan
Ethiopia
Somalia
Kenya
Gabon
DR Congo
Tanzania
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
South Africa
Gulf of Guinea
Gulf of Aden
Arabian Sea
Bay of Bengal
Laccadive Sea
Myanmar (Burma)
Thailand
Vietnam
South China Sea
Philippines
Gulf of Thailand
Malaysia
Indonesia
Banda Sea
Arafura Sea
Australia
Great Australian Bight
Guatemala
Nicaragua
Venezuela
Guyana
Suriname
Colombia
Ecuador
Peru
Brazil
Bolivia
Paraguay
Chile
Argentina
Uruguay

Event details:

- New state
- Same country
- Same device

● Not suspicious

● Suspicious

✸ Very suspicious

Canada
Hudson Bay
AB
MB
BC
SK
ON
QC
NL
NB
PE
NS
WA
MT
ND
MN
OR
ID
SD
WY
WI
MI
ME
NH
NY
MA
NE
IA
IL
IN
PA
NV
UT
United States
CO
KS
MO
MD
DE
KY
WV
VA
CA
OK
AR
TN
NC
AZ
NM
MS
AL
GA
SC
TX
LA
FL
Gulf of Mexico
Mexico
Cuba
Puerto Rico
Guatemala
Caribbean Sea
Nicaragua
Venezuela
Guyana
Suriname
Colombia
RR
AP
Ecuador
AM
PA
MA
CE
RN
PB
PI
PE
AL
SE
AC
Brazil
Peru
RO
TO
MT
BA
GO
Bolivia
MG
ES
MS
SP
RJ
Paraguay
PR
SC
Chile
RS
Uruguay
Argentina
North Atlantic Ocean
South Atlantic Ocean
Norway
Finland
Russia
United Kingdom
Denmark
Ireland
Belarus
Poland
Germany
Ukraine
Austria
France
Romania
Italy
Spain
Portugal
Greece
Turkey
Kazakhstan
Mongolia
Uzbekistan
Kyrgyzstan
Turkmenistan
Sea of Japan
South Korea
Jap
China
Tunisia
Syria
Iraq
Iran
Afghanistan
Morocco
Pakistan
East China Sea
Nepal
Algeria
Libya
Egypt
Western Sahara
Saudi Arabia
India
Oman
Myanmar (Burma)
Mauritania
Mali
Niger
Chad
Sudan
Yemen
Thailand
South China Sea
Burkina Faso
Gulf of Aden
Arabian Sea
Bay of Bengal
Vietnam
Philippines
Guinea
Nigeria
Ghana
Ethiopia
South Sudan
Gulf of Thailand
Laccadive Sea
Gulf of Guinea
Malaysia
Somalia
Kenya
Gabon
DR Congo
Indonesia
Tanzania
Banda Sea
Arafura Sea
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
Indian Ocean
NT
Australia
WA
SA
South Africa
Great Australian Bight

Event details:

- New country
- Same device

● Not suspicious

● Suspicious

✸ Very suspicious

Canada
Hudson Bay
AB
MB
BC
SK
ON
QC
NL
NB
PE
NS
WA
MT
ND
MN
OR
ID
SD
WY
WI
MI
NE
IA
IL
IN
NY
ME
NH
MA
PA
MD
DE
NV
UT
United States
CO
KS
MO
KY
WV
VA
CA
AZ
NM
OK
AR
TN
NC
SC
MS
AL
GA
TX
LA
FL
Gulf of Mexico
Mexico
Cuba
Puerto Rico
Guatemala
Caribbean Sea
Nicaragua
Venezuela
Guyana
Suriname
Colombia
RR
AP
Ecuador
AM
PA
MA
CE
RN
PB
PI
PE
AL
SE
AC
Brazil
Peru
RO
MT
TO
BA
GO
Bolivia
MG
ES
MS
SP
RJ
Paraguay
PR
SC
Chile
RS
Uruguay
Argentina
North Atlantic Ocean
South Atlantic Ocean
Norway
Finland
Russia
Denmark
United Kingdom
Ireland
Belarus
Poland
Germany
Ukraine
Austria
France
Romania
Italy
Spain
Portugal
Greece
Turkey
Kazakhstan
Mongolia
Uzbekistan
Kyrgyzstan
Turkmenistan
Sea of Japan
South Korea
China
Japan
Syria
Iraq
Iran
Afghanistan
Pakistan
Tunisia
Morocco
Algeria
Libya
Egypt
Western Sahara
Saudi Arabia
Oman
Nepal
India
East China Sea
Myanmar (Burma)
Mauritania
Mali
Niger
Chad
Sudan
Yemen
Gulf of Aden
Arabian Sea
Bay of Bengal
Thailand
South China Sea
Vietnam
Philippines
Burkina Faso
Guinea
Ghana
Nigeria
South Sudan
Ethiopia
Gulf of Guinea
Laccadive Sea
Gulf of Thailand
Malaysia
Somalia
Kenya
Gabon
DR Congo
Indonesia
Banda Sea
Tanzania
Arafura Sea
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
Indian Ocean
NT
Australia
WA
SA
South Africa
Great Australian Bight

Event details:

- New country
- IP from data center
- New device
- Unusual event timing

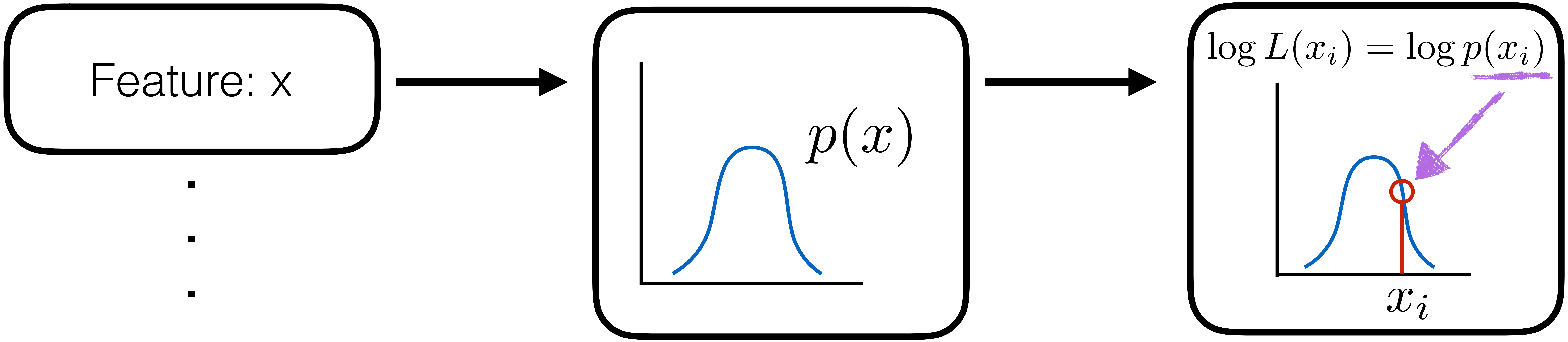
● Not suspicious

● Suspicious

✸ Very suspicious

Unsupervised anomaly detection

All features | **Calculate feature distribution** | **Likelihood of observed feature value**



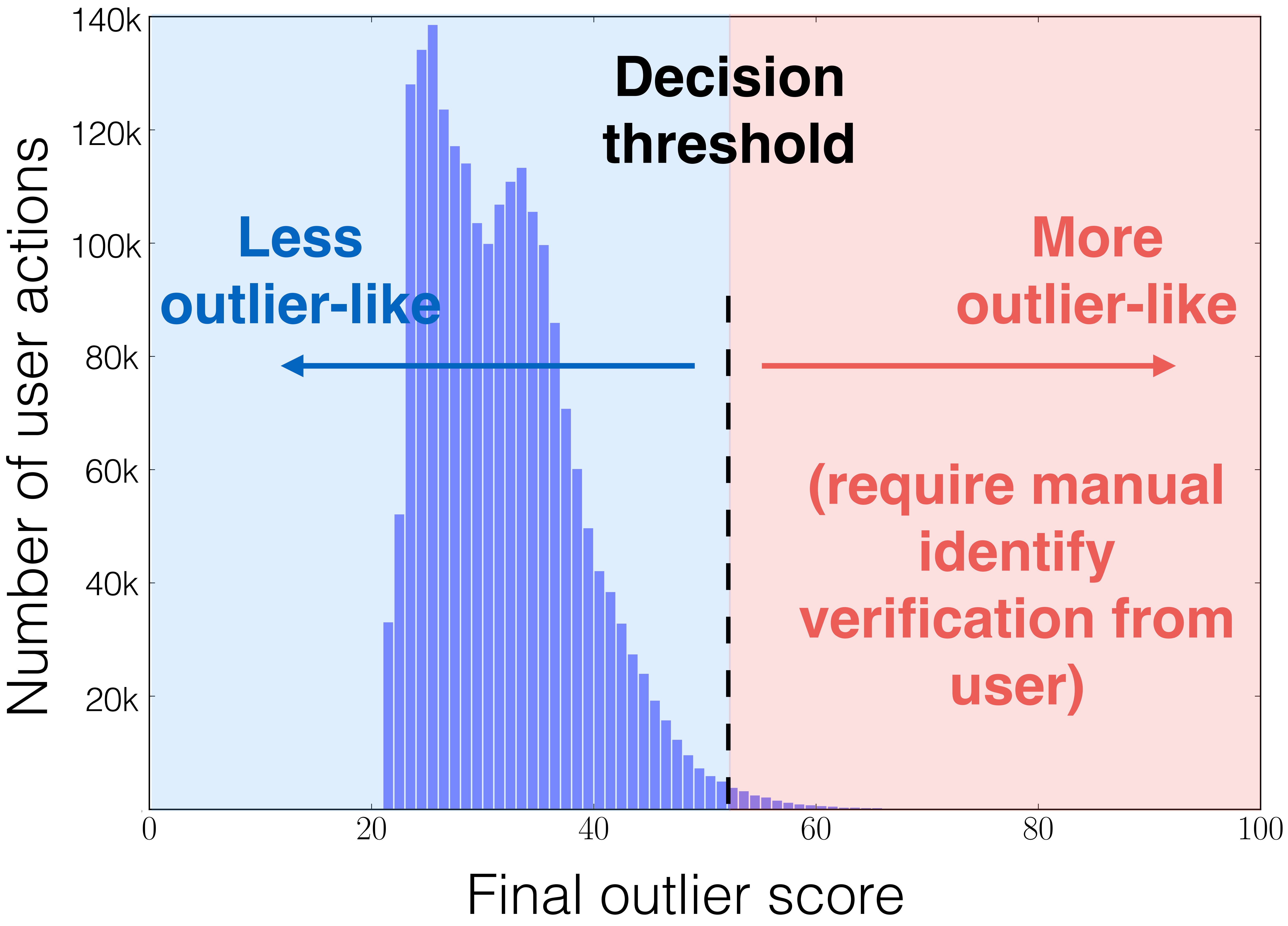
For event i, combine log-likelihood over all features:

$$\log L_i = \log L(x_i) + \log L(y_i) + \ldots$$

Single final score for outlier detection

Tune decision threshold

- Can **validate** using list of known compromised accounts
- Area under ROC curve = **0.95**
- For this choice of threshold: **79% recall** with **5% false positive rate**



Product delivers measurable improvement

- **Improved** the recall of compromised accounts by >2x compared to the baseline model with the same false positive rate
- **Fast** to train and use
- **Interpretable** feature importance
- **Extendable** to include new features
- **More details** available at jverbus.github.io

James Verbus