PRICE MY BIKE!

QUICKLY PRICING YOUR USED ITEMS

BRUNO AFONSO, INSIGHT FELLOW

USED AND SECOND-HAND GOODS INDUSTRY IS BOOMING

- Environmental concerns have driven people to the used market
- Declutter and save money
- Canadian second-hand economy was 28.5B CAD (2018)
- Most Canadians take part in the second-hand economy
  - 82% (2016)
  - 85% (2018)

PRICING USED ITEMS IS HARD… CAN WE SUGGEST A PRICE?

If only I could quickly price the stuff I want to "declutter" …

4 Price

Suggested price: 200

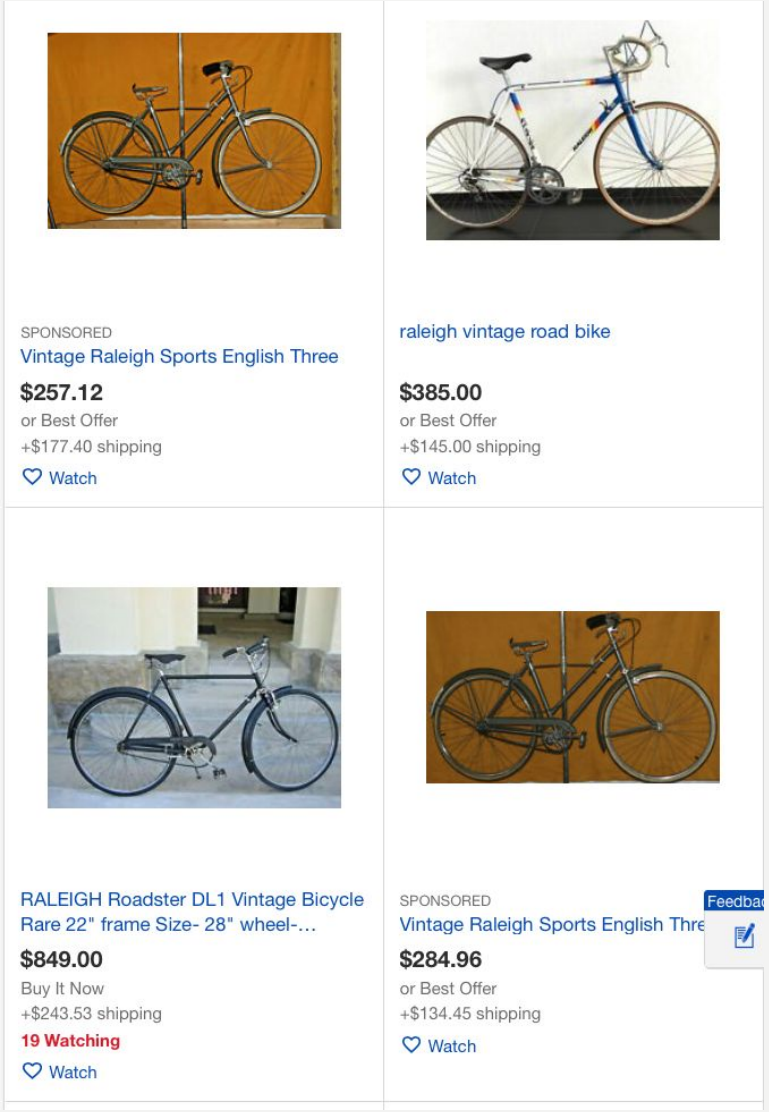
Price: $

Free

Please Contact

Swap / Trade

EXAMPLE MARKET: USED BIKES



- Bicycles are hard to price given the the high number of components and wide range of quality
- Bike components can vary one order of magnitude.
- Wheels: from $50 to $2000
- Depreciation is brand and model dependent
- Prices change based on trends

DATA SOURCE: EBAY

Data acquisition

```
'Title': 'Specialized Hard Rock
Vintage Mountain Bike X Large 21"
''MTB Suntour XC
Charity!','ItemSpecifics':
{'NameValueList':
[{'Name': 'Wheel Size',
'Value': '26"'},
{'Name': 'Model',
'Value': 'Hard Rock'},
{'Name': 'Suspension',
'Value': 'No Suspension'},
{'Name': 'Handlebar Type',
'Value': 'Flat Bar'},
{'Name': 'Number of Gears',
'Value': '21'},
{'Name': 'Frame Size',
'Value': '53cm (21")'},
{'Name': 'Type',
'Value': 'Mountain Bike'},
{'Name': 'Features',
'Value': ['Rear Rack',
'Adjustable Seat']},
{'Name': 'Color', 'Value':
'Red'},
{'Name': 'Frame Material',
'Value': 'Steel'},
{'Name': 'Brand',
'Value': 'Specialized'},
{'Name': 'Brake Type',
'Value': 'Cantilever'},
{'Name': 'Manufacturer Color',
'Value': 'Fire Engine Red'},
```


eBay Bicycles category

- 14k sold items (Backlog of 90 days)
- **final** transaction price
- Manually entered information by sellers

- Brand
- Color
- Material
- …

Data acquisition

Data cleaning

{'AutoPay': 'true',
 'BestOfferEnabled': 'false',
 'BidCount': '27',
 'BuyItNowAvailable': 'true',
 'BuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConditionDisplayName': 'Used',
 'ConditionID': '3000',
 'ConvertedBuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConvertedCurrentPrice':
{'_currencyID': 'USD', 'value':
'69.0'},
 'Country': 'US',

112 item specific categories

↓ Pre-process the text and remove low represented categories (<1%)

16 item specific categories

↓ Further prune category entries and replace missing values with “other”

4 categories

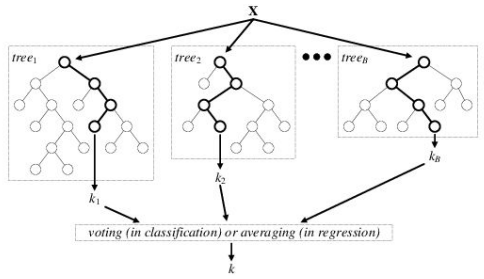
DATA PIPELINE

Data acquisition

Data cleaning

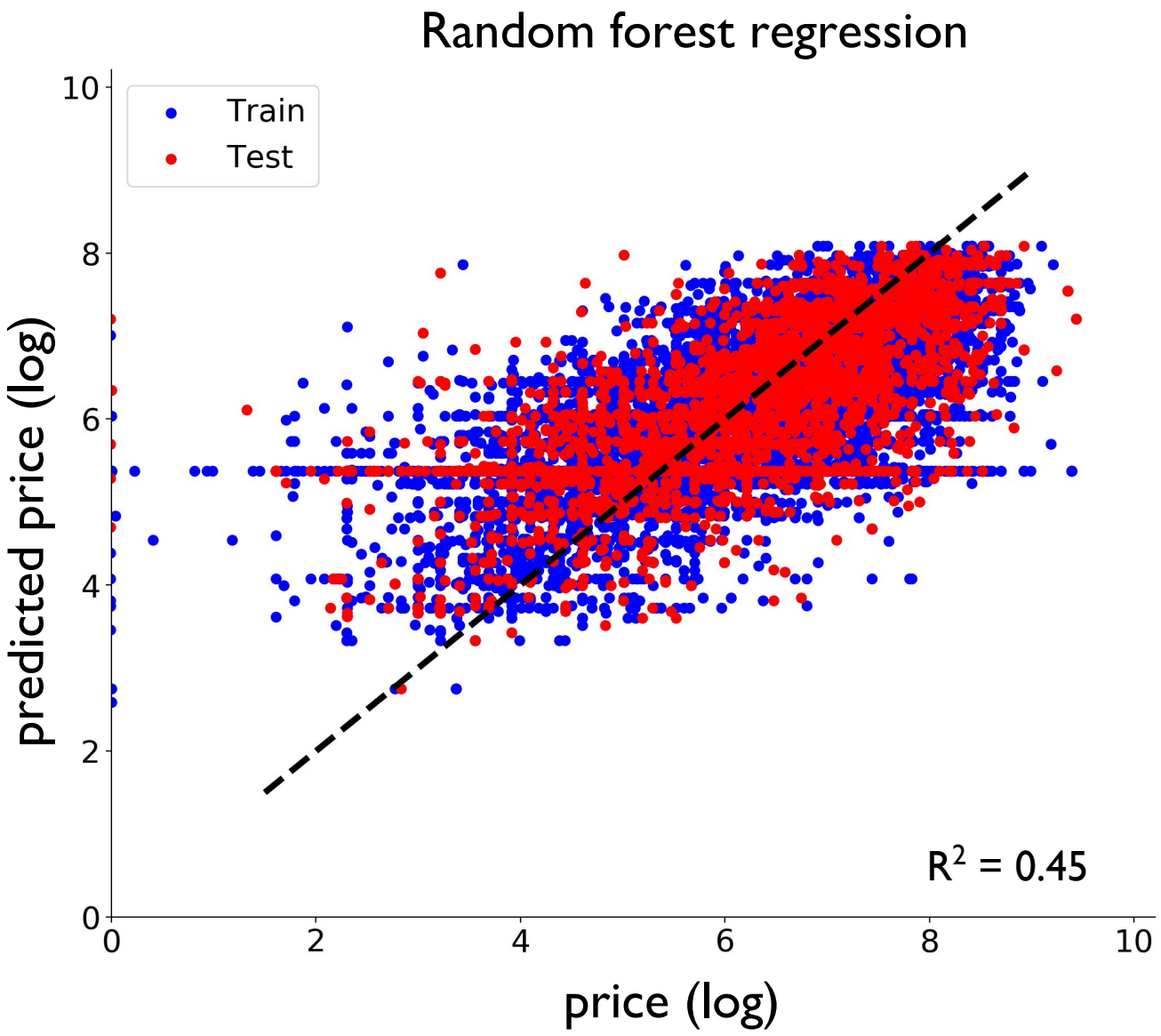
Model building

{'AutoPay': 'true',
 'BestOfferEnabled': 'false',
 'BidCount': '27',
 'BuyItNowAvailable': 'true',
 'BuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConditionDisplayName': 'Used',
 'ConditionID': '3000',
 'ConvertedBuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConvertedCurrentPrice':
{'_currencyID': 'USD', 'value':
'69.0'},
 'Country': 'US',



Random forest regressor

PRICING MODEL



- Random Forest Regression
  - 4 categorical features
  - 14k original entries, (~3k with no missing values)
  - Training 80%
  - Test 20%
  - Replaced all missing values for “other”
  - R^2 = 0.45

DATA PIPELINE

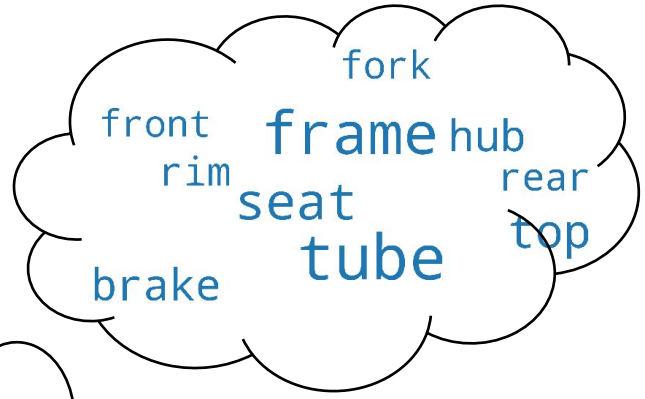
Data acquisition

Data cleaning

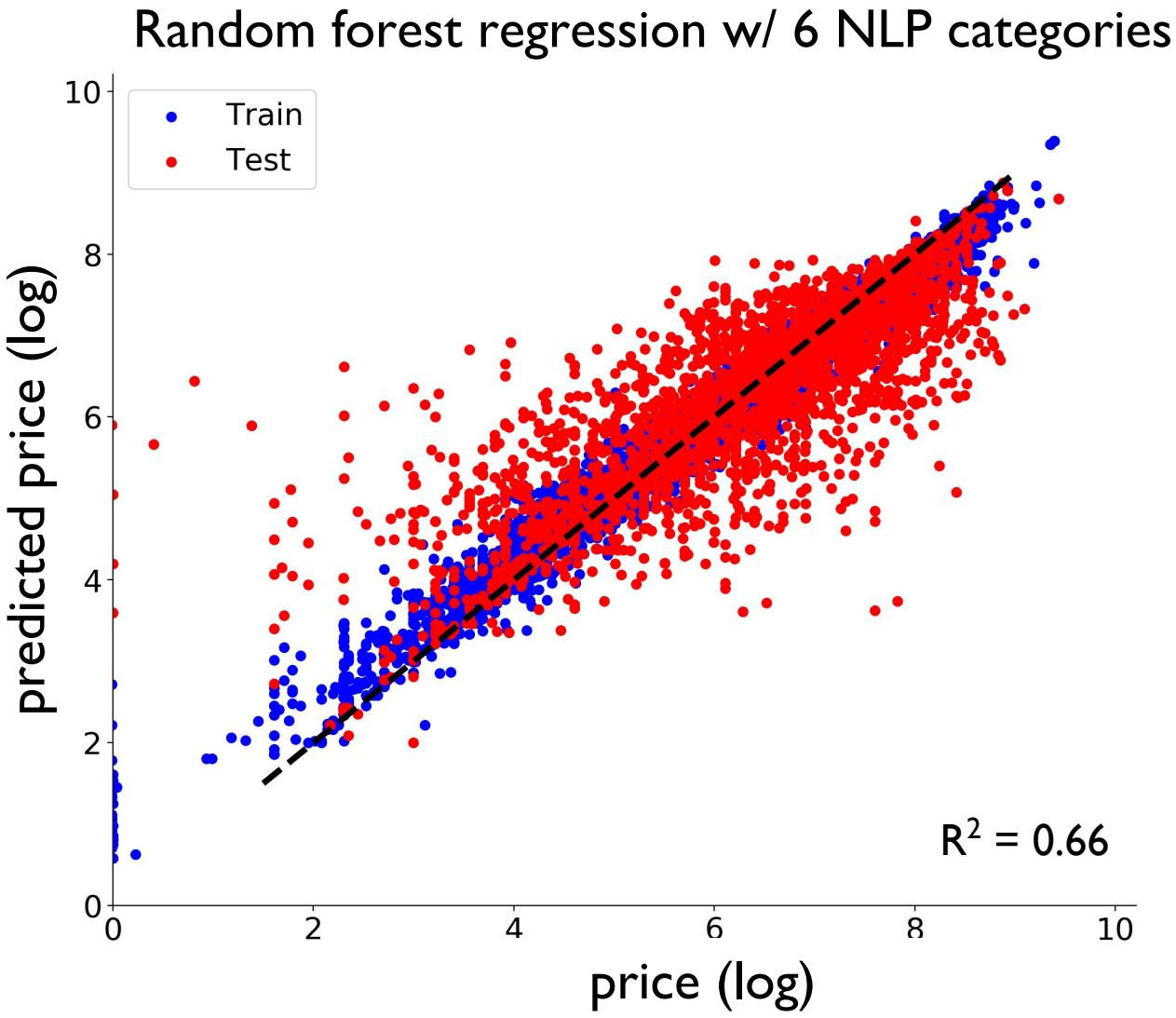
Feature engineering

Topic modeling

{'AutoPay': 'true',
 'BestOfferEnabled': 'false',
 'BidCount': '27',
 'BuyItNowAvailable': 'true',
 'BuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConditionDisplayName': 'Used',
 'ConditionID': '3000',
 'ConvertedBuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConvertedCurrentPrice':
{'_currencyID': 'USD', 'value':
'69.0'},
 'Country': 'US',



PRICING MODEL WITH NLP



- Random Forest Regression with NLP
  - 4 categorical features
  - 14k original entries, (~3k with no missing values)
  - Training 80%
  - Test 20%
  - 6 new features based on topic modeling (NLP)
  - R^2 = 0.66

DATA PIPELINE

Data acquisition

Data cleaning

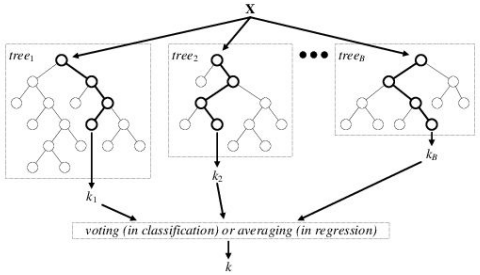
Feature engineering

Model building

Web app

Topic modeling

{'AutoPay': 'true',
 'BestOfferEnabled': 'false',
 'BidCount': '27',
 'BuyItNowAvailable': 'true',
 'BuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConditionDisplayName': 'Used',
 'ConditionID': '3000',
 'ConvertedBuyItNowPrice':
{'_currencyID': 'USD', 'value':
'149.99'},
 'ConvertedCurrentPrice':
{'_currencyID': 'USD', 'value':
'69.0'},
 'Country': 'US',



Random forest regressor

type cruiser

wheelsize 20"

framematerial carbonfiber

brand bmc

Description

shimano groups but is dirty

DEMO !

Price my Bike demo time !

type cruiser

wheelsize 20"

framematerial carbonfiber

brand bmc

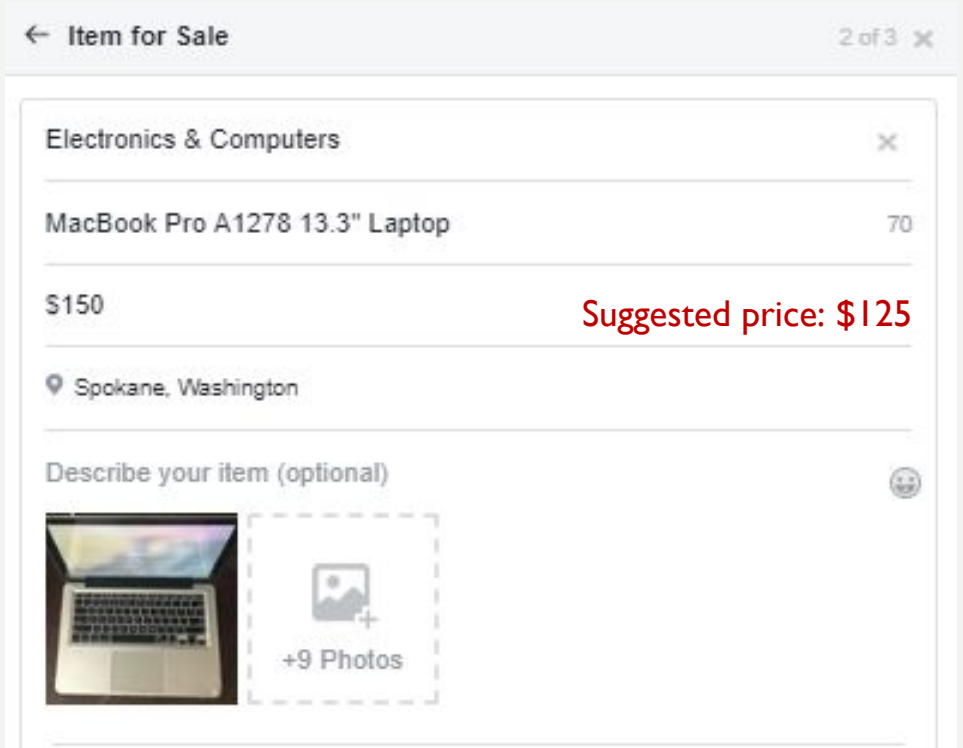
Description

shimano groupset but is dirty

Estimated price is $543

DEPLOYMENT OPTIONS

- Create a website for end users selling as well as recommendation website for users looking for items
- Build an API for Kijiji, craigslist, eBay, Facebook
  - aggregate data from all these companies for custom tailored estimations

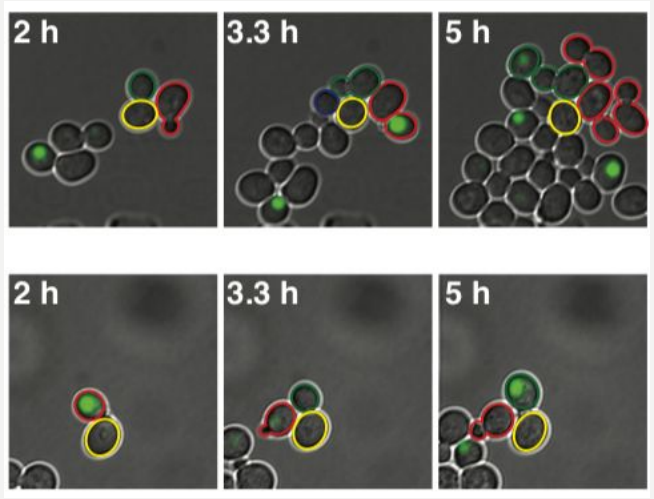


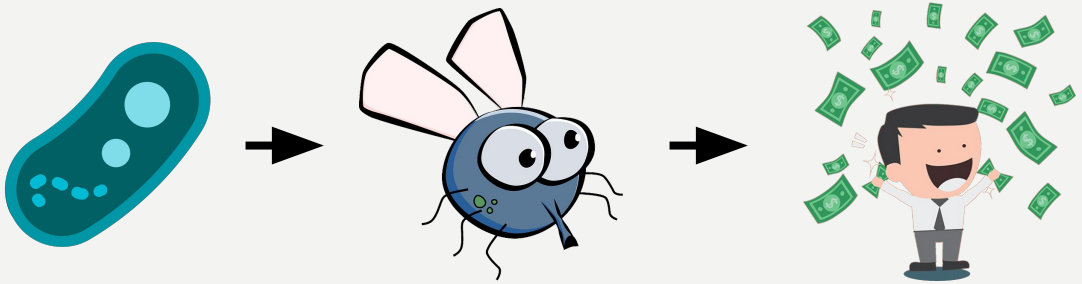
BRUNO AFONSO

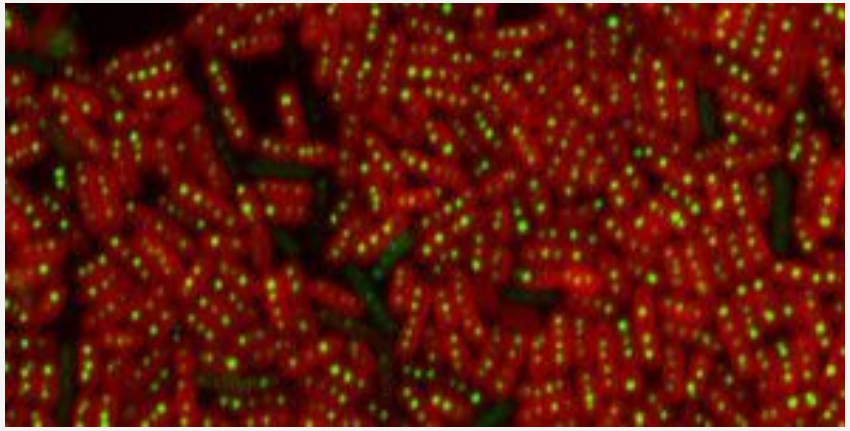
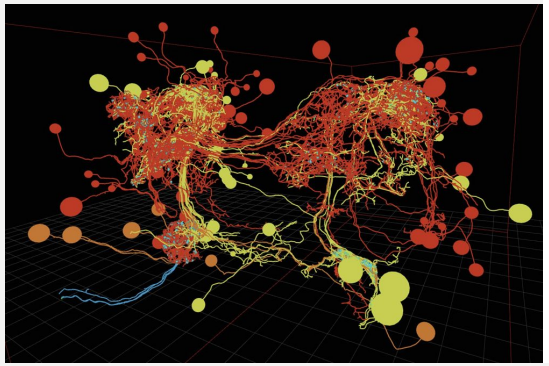
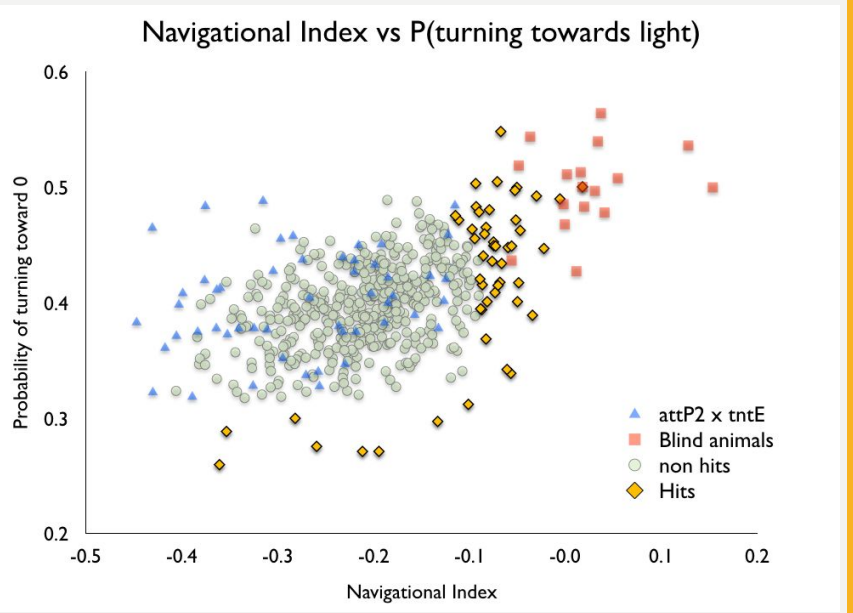
I like data

I love noisy data

HARVARD MEDICAL SCHOOL | BLAVATNIK INSTITUTE SYSTEMS BIOLOGY







hhmi | janelia Research Campus