Machine Learning Driven Social Engineering

Charles Herring
Co-Founder, CTO

me@CharlesHerring.com
https://CharlesHerring.com
@charlesherring

About Charles

1995-2002: Forward Deployed US Navy Hornet Avionics Tech

2002-2005 US Naval Postgraduate School Network Security Group Division Officer
sk3wl 0f r00t team member

2003-2008: InfoWorld Test Center
Contributing Product Reviewer – Network and Information Security

2005-2012 DoD Security, Data & Workflow Consultant

2012-2016: Consulting Security Architect for Lancope then Cisco Systems

2016 CTO & co-founder at WitFoo

Disclaimer

Information in this presentation is intended to protect against illegal or harmful social engineering and human manipulation or to be used in benevolent endeavors.

Please do good and obey the law.

Old School Social Engineering

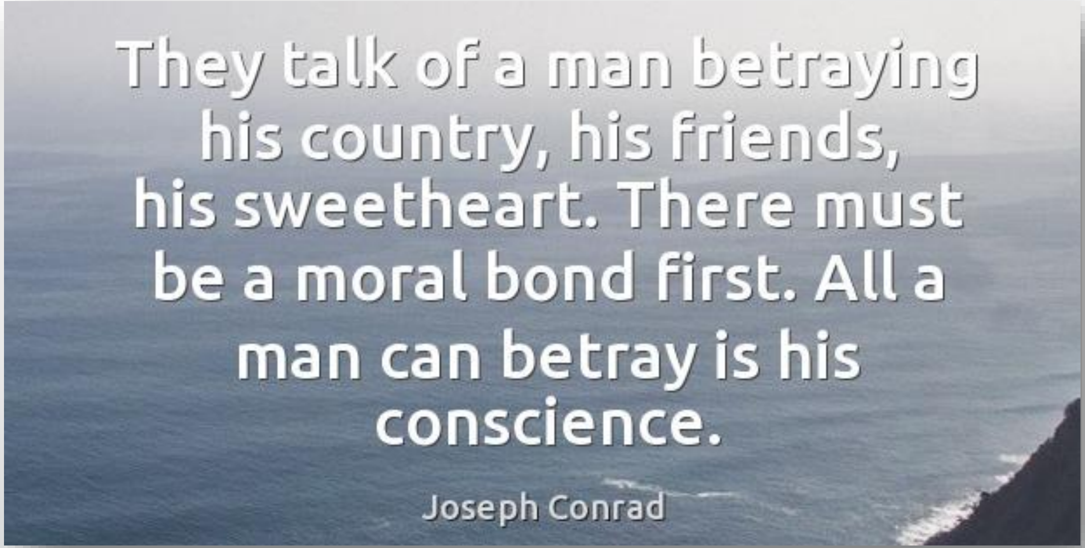
- Based on “Confidence Games” (Cons)
- Target a “mark” to achieve inappropriate action
- Requires understanding of “mark” (agent)
- Manipulates existing human vulnerability

Second Generation Social Engineering

- Bulk, mass attacks
- Phishing
- Robocalls

Social Engineering with ML

- Need the Cohort (not a specific agent) to take action
- Long term training will transform agent values (create vulnerability)

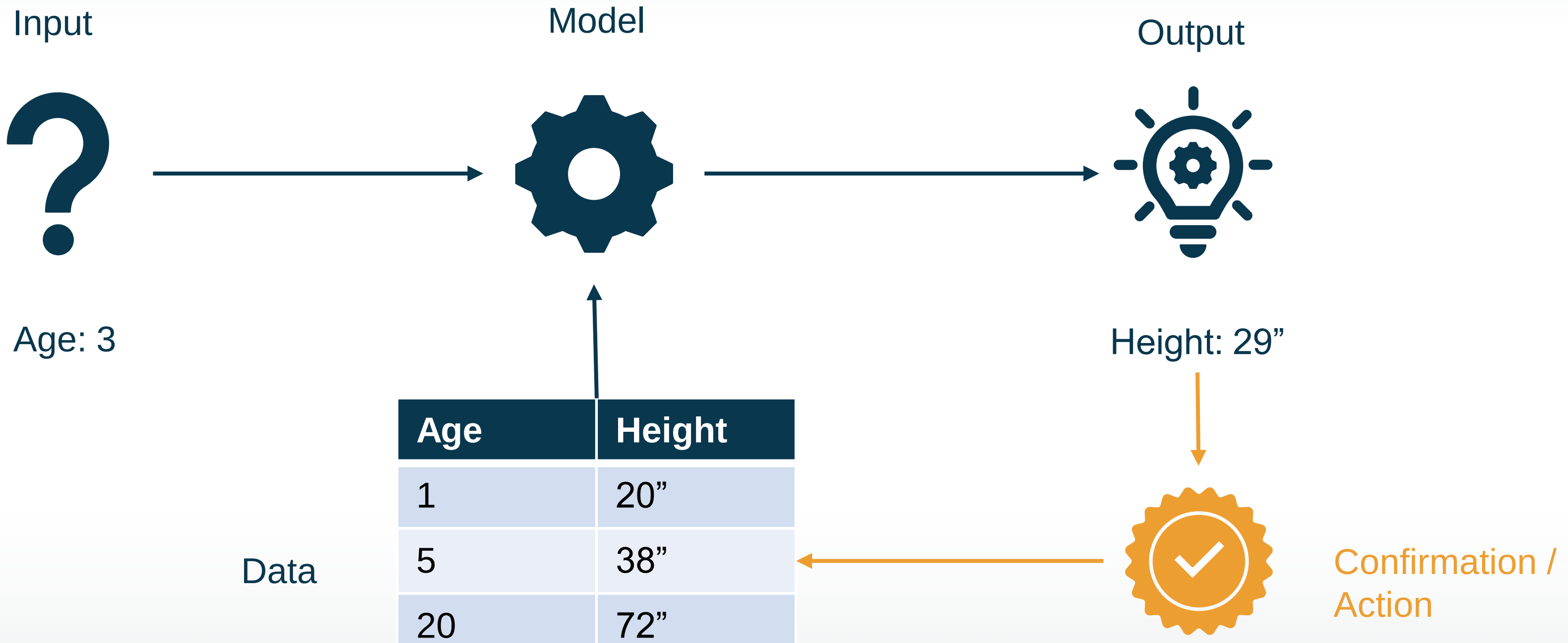


Relentless Manipulation

- *"We've created a world in which online connection has become primary. Especially for younger generations. And yet, in that world, anytime two people connect, the only way it's financed is through a sneaky third person who's paying to manipulate those two people. So, we've created an entire global generation of people who were raised within a context with the very meaning of communication, the very meaning of culture, is manipulation."* – Jaron Lainer in *The Social Dilemma*

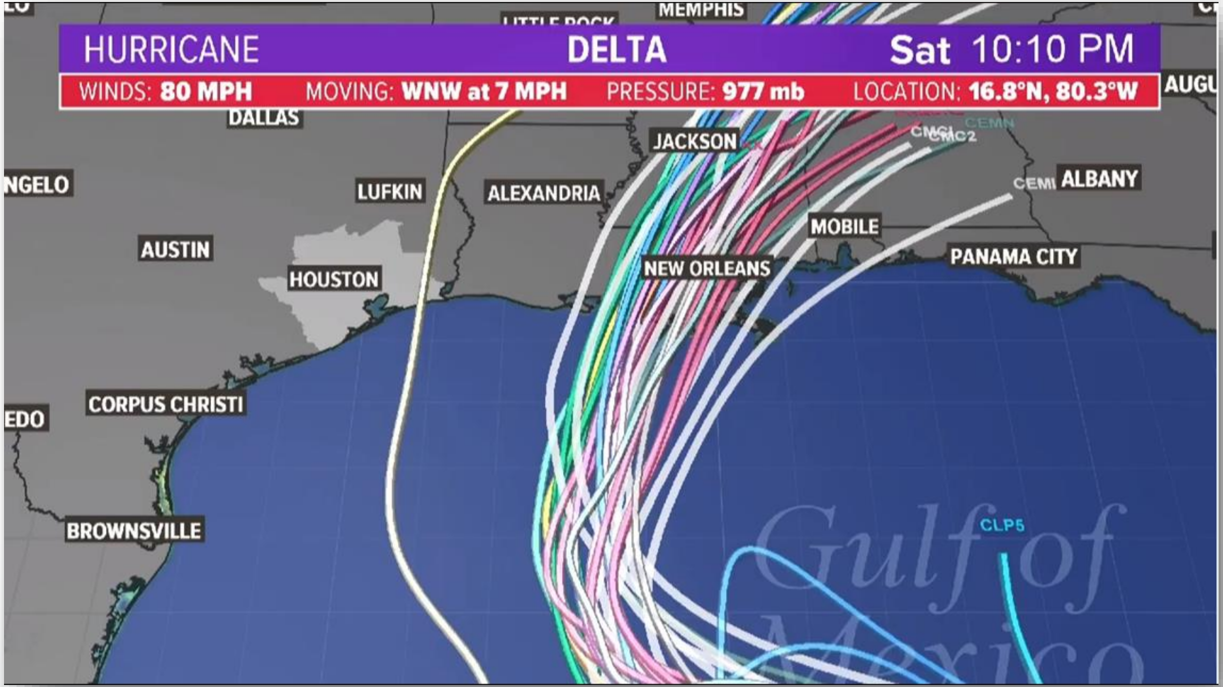
Machine Learning Basics

Supervised Learning



Age	Height
1	20”
5	38”
20	72”
50	72”

Data and Models



witfoo

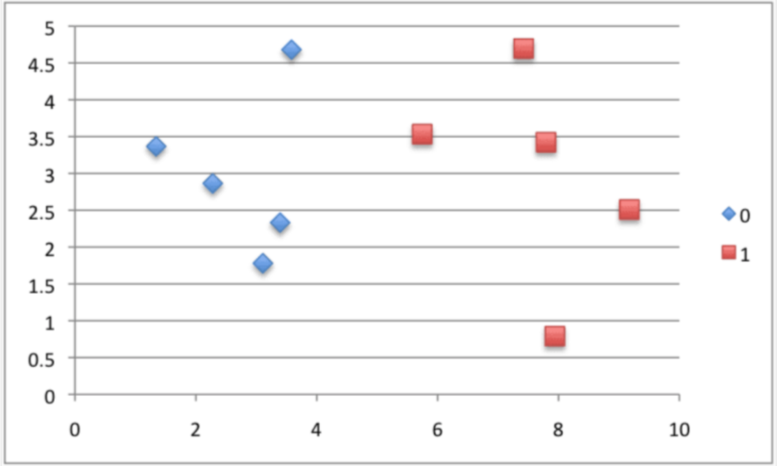
Common Models

- ***Naive Bayes***
- Nearest Neighbor
- Decision Trees
- Linear Regression
- Support Vector Machines (SVM)
- Neural Networks

https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861

Naïve Bayesian Python

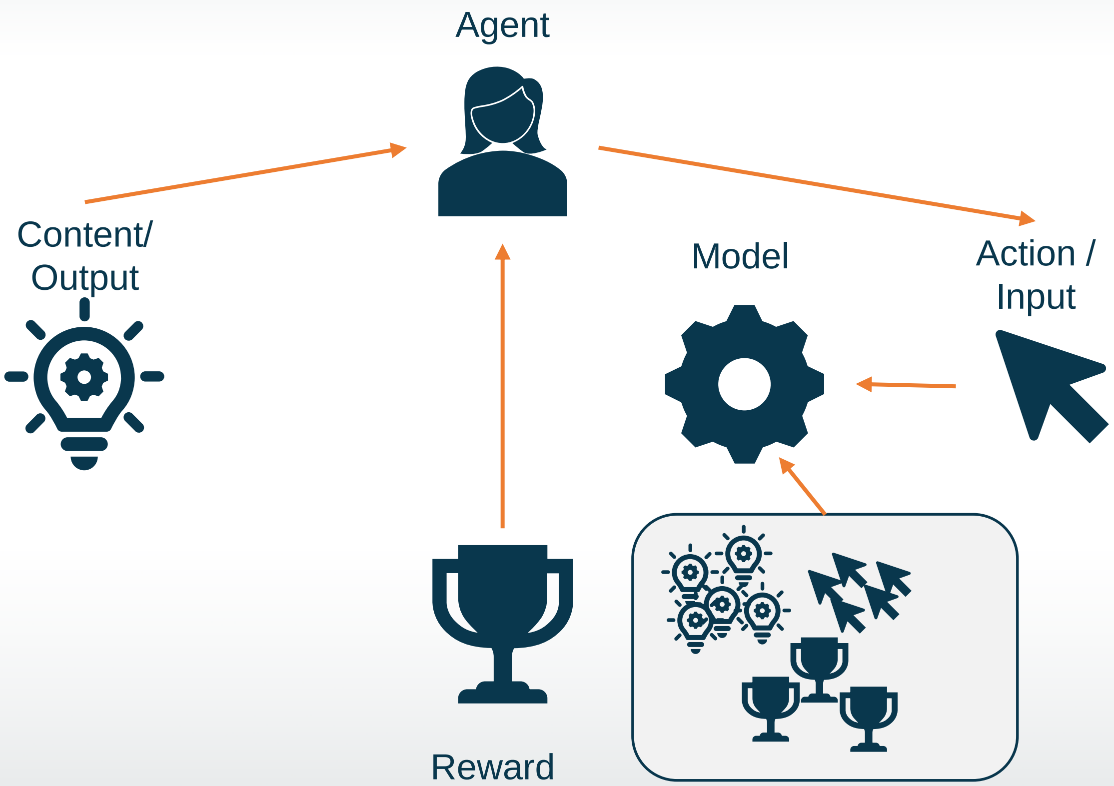
- https://machinelearningmastery.com/naive-bayes-classifier-scratch-python/



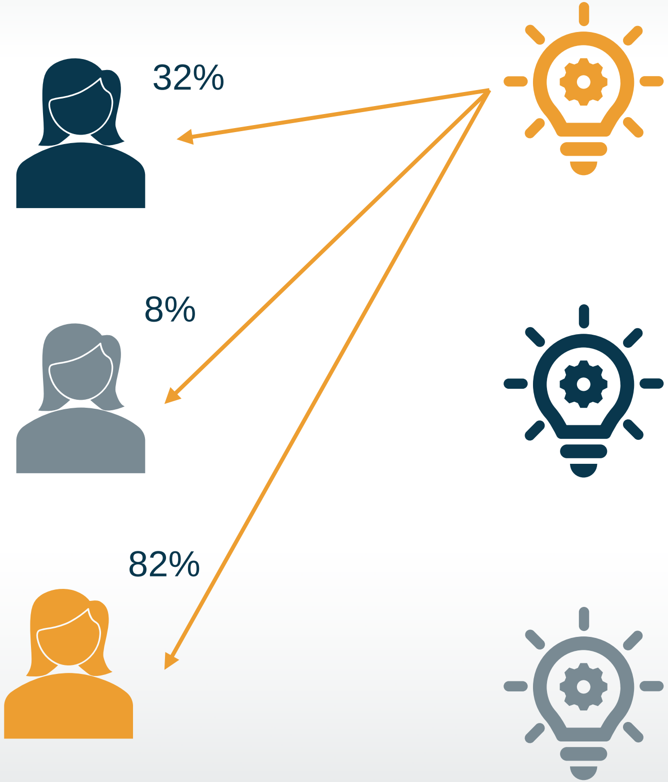
Scatter Plot of Small Contrived Dataset for Testing the Naive Bayes Algorithm

```python
# Example of separating data by class value

# Split the dataset by class values, returns a dictionary
def separate_by_class(dataset):
    separated = dict()
    for i in range(len(dataset)):
        vector = dataset[i]
        class_value = vector[-1]
        if (class_value not in separated):
            separated[class_value] = list()
        separated[class_value].append(vector)
    return separated

# Test separating data by class
dataset = [[3.393533211,2.331273381,0],
    [3.110073483,1.781539638,0],
    [1.343808831,3.368360954,0],
    [3.582294042,4.67917911,0],
    [2.280362439,2.866990263,0],
    [7.423436942,4.696522875,1],
    [5.745051997,3.533989803,1],
    [9.172168622,2.511101045,1],
    [7.792783481,3.424088941,1],
    [7.939820817,0.791637231,1]]
separated = separate_by_class(dataset)
for label in separated:
    print(label)
    for row in separated[label]:
        print(row)
```

Reinforcement Learning



A/B Testing of Output / Signal / Content



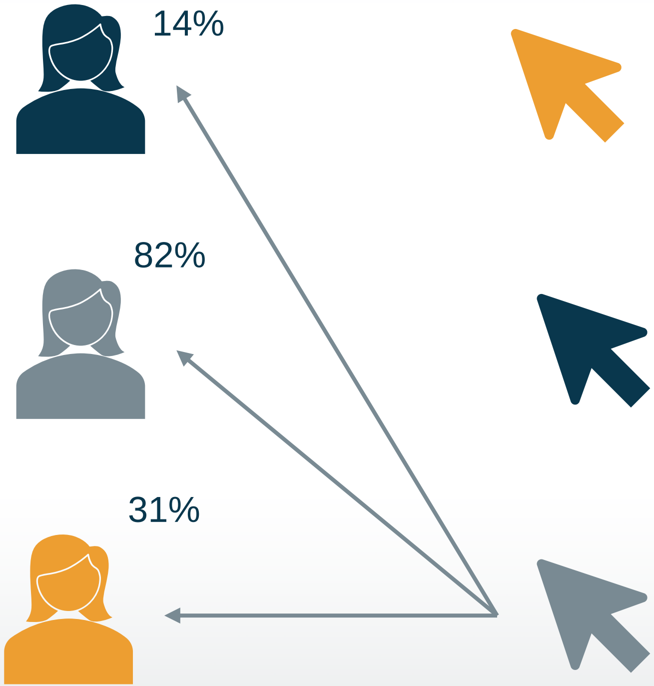
Types of Content

- Text Ad
- Video Ad
- Promoted Pages
- Email
- Snail Mail
- Images

Message Considerations

- Alignment with Values
- Suspicion Triggers
- Language & Dialect Use
- Use of Nostalgia
- Trauma Triggers (Crossing a line)
- Bias Reinforcement
- Trusted Corroboration
- Social Corroboration
- Pain vs. Reward

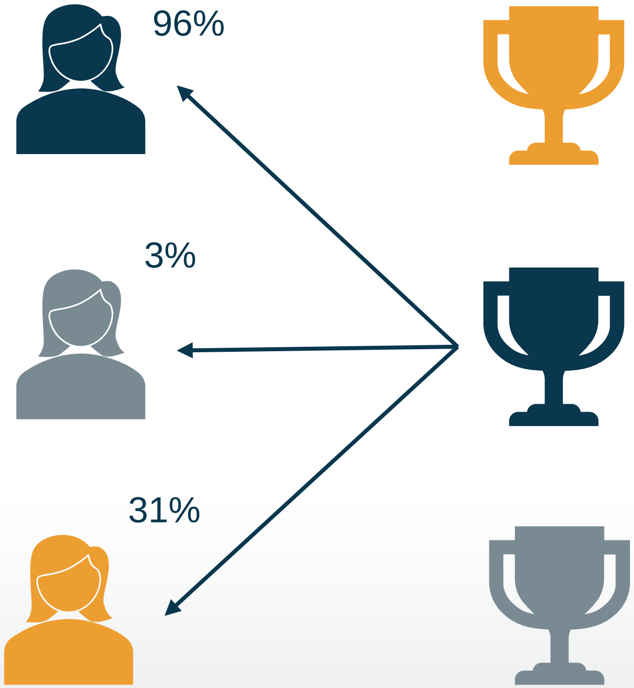
A/B Testing of Action



Types of Actions

- Click to Page
- Watch Video
- Play Game
- Fill out Form
- Join a Group
- Install an Application
- Comment / Engage

A/B Testing of Reward



Types of Reward

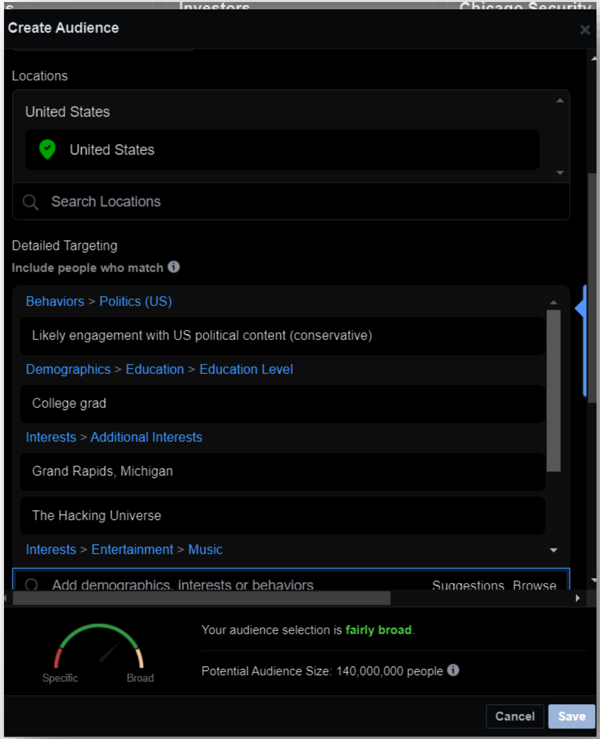
- Public Praise – like, retweet, comment
- Artificial Prize - badge, points
- Physical Prize - cash, gift card, prize, beer, meal
- Education Credit – CEU, certification
- Celebrity Interaction – virtual or physical
- Promise of job
- Romantic Relationship or Encounter
- Confirmation Reinforcement

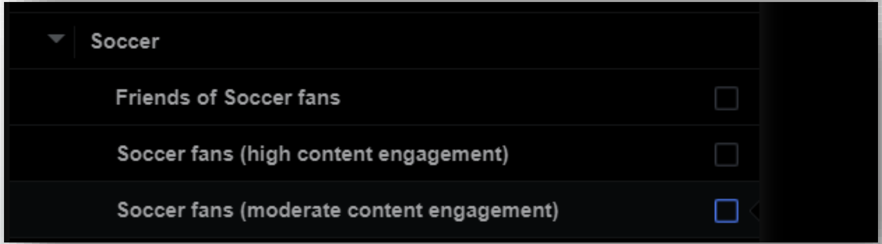
Cohort Targeting

Common Cohort Dimensions

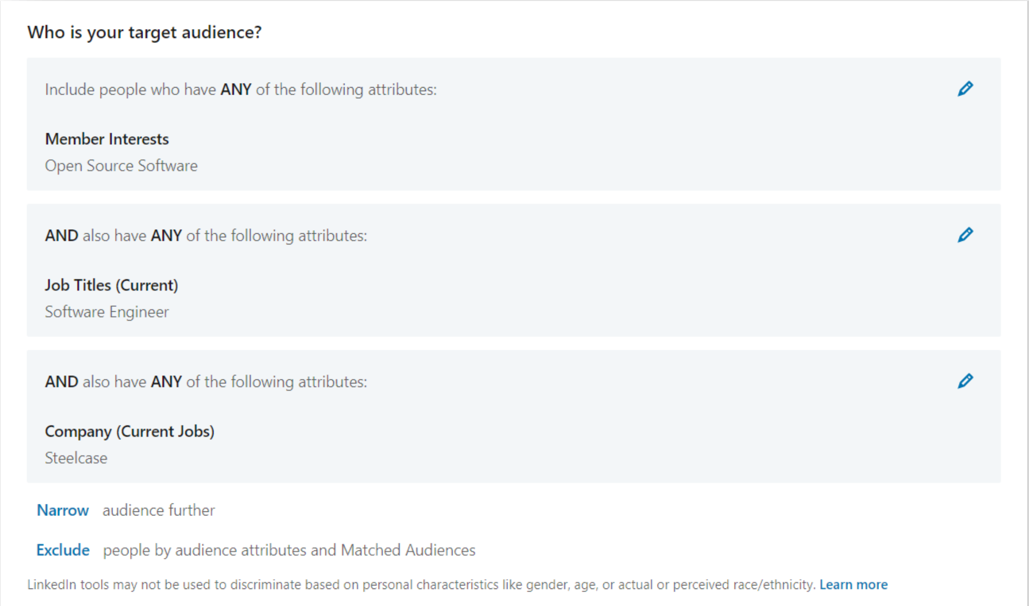
- Gender
- Age
- Ethnicity
- Religious Belief
- Income
- Location
- Family Status
- Job Type
- Employer Type
- Education Level
- Affiliations
- Relationships
- Entertainment
- Credit Rating
- Net Worth
- Political Association

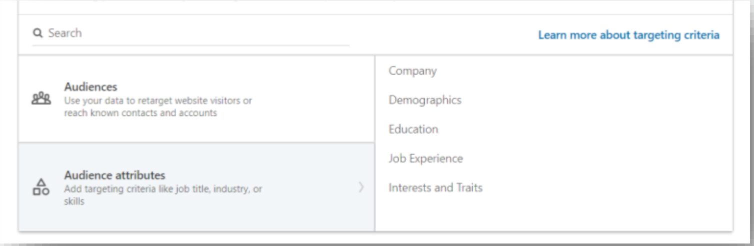
Facebook Audiences





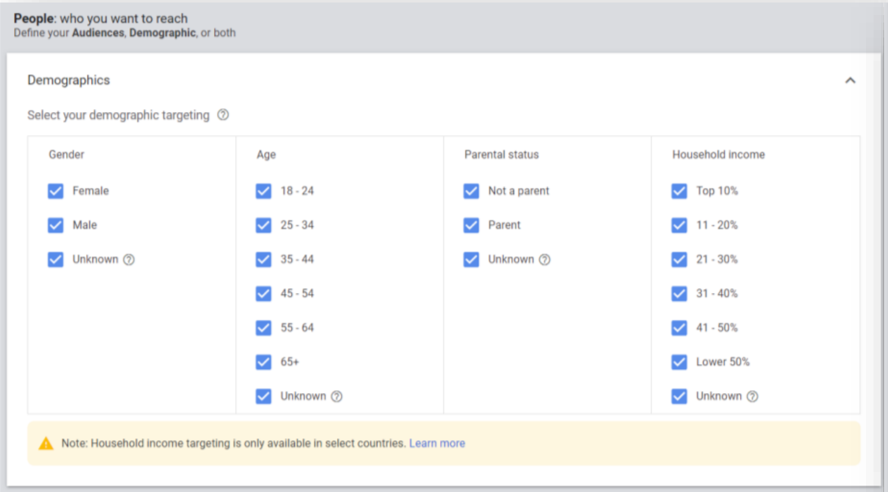
LinkedIn Audience

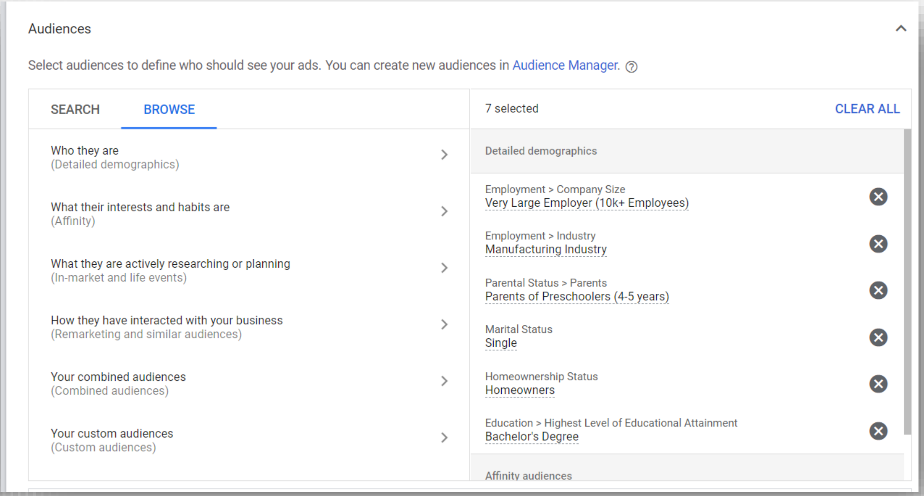




Useable on Bing Search

Google Audience



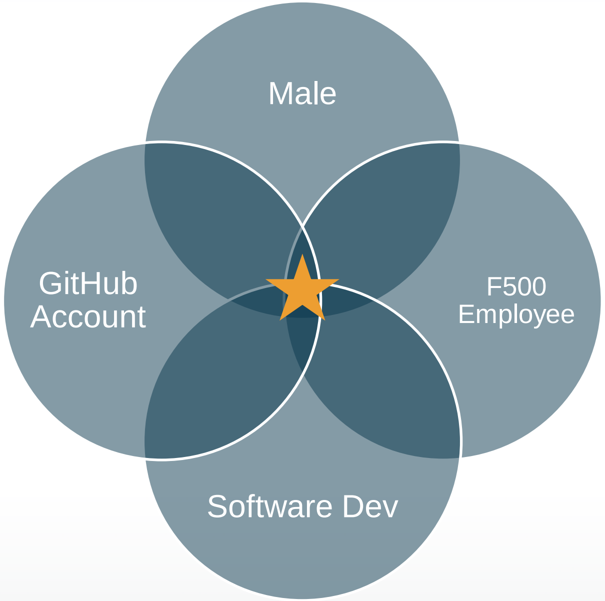


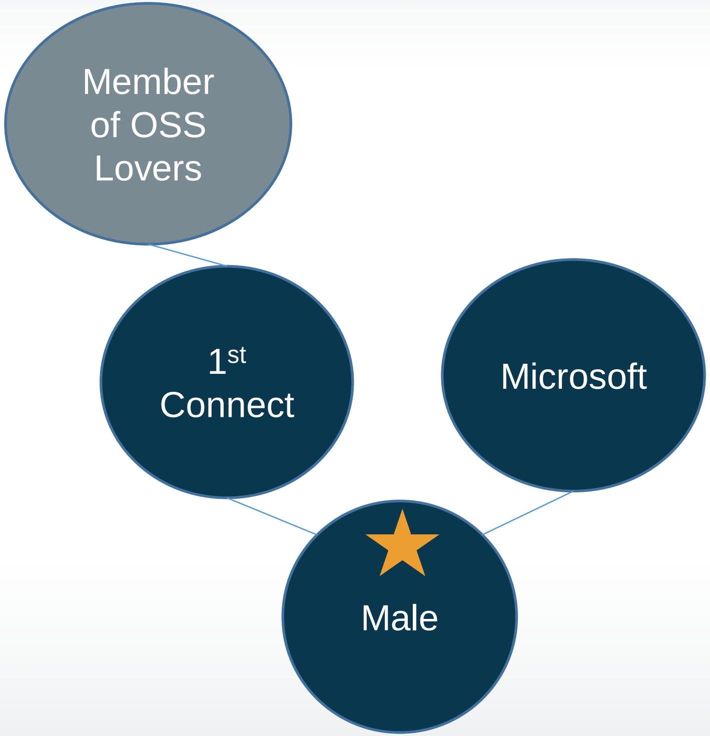
witfoo

Other Sources of Information

- Breach Data (Cambridge Analytica, etc)
- Intellius
- Scraping Social Media Sites
- Public Records

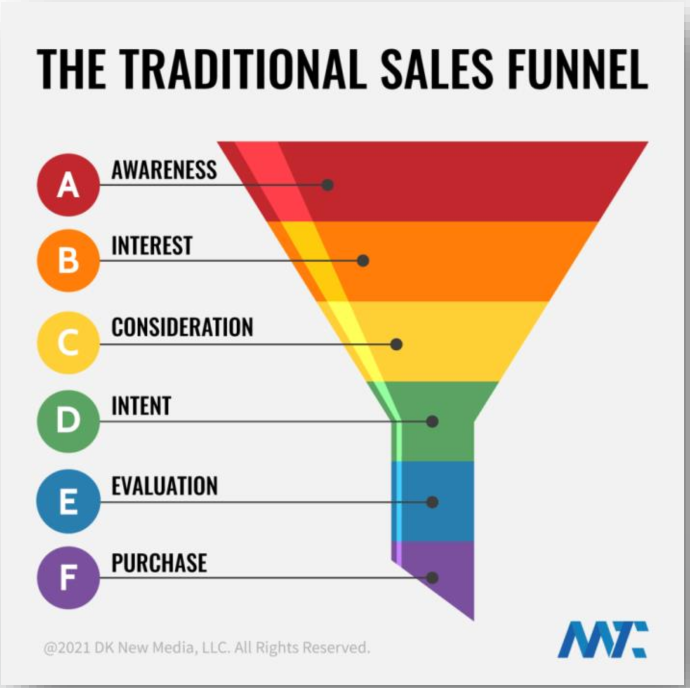
Cohort Collison vs Relationship





Launching Campaigns

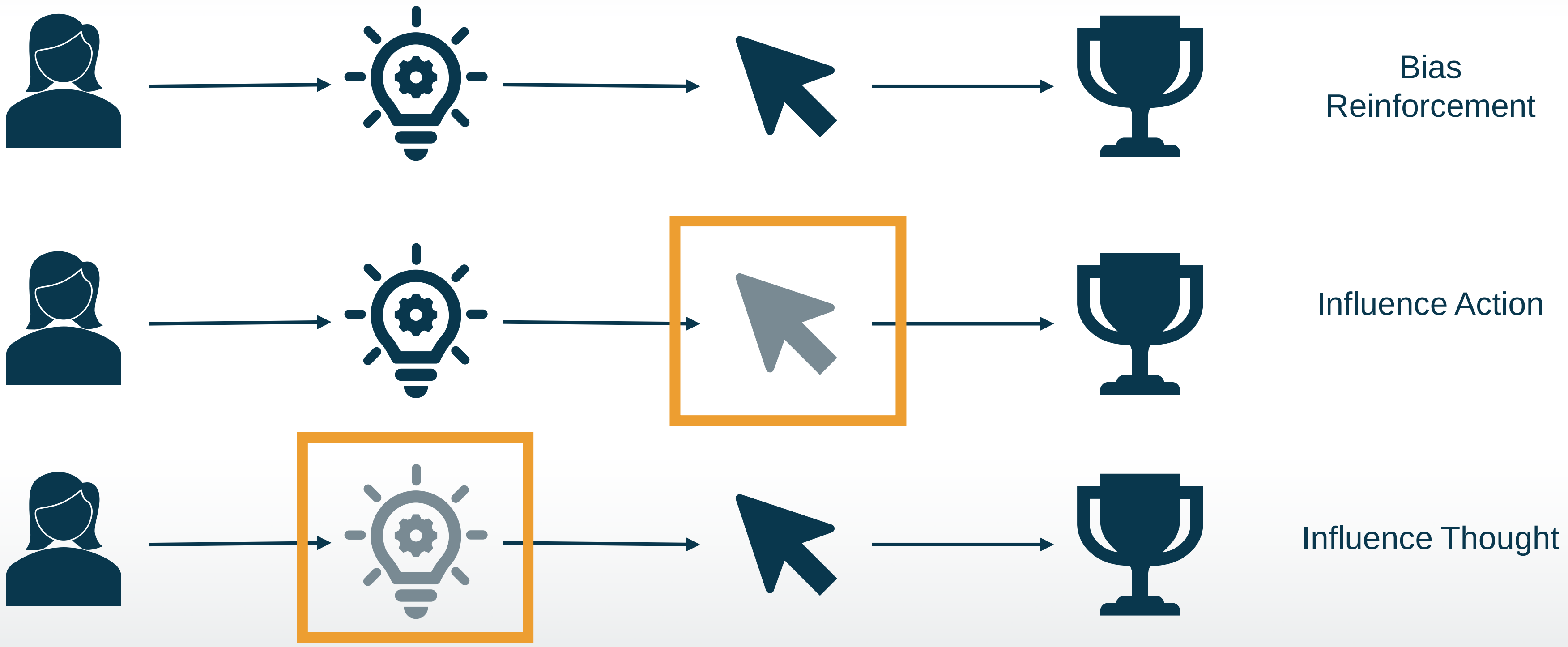
Humans Move Slowly



Sales cycles as detailed here can take **months or years** to complete as agents progress through the process.

Successful progressions will have small, incremental steps requiring the lowest amount of agent change per stage.

Agent Transformation



witfoo

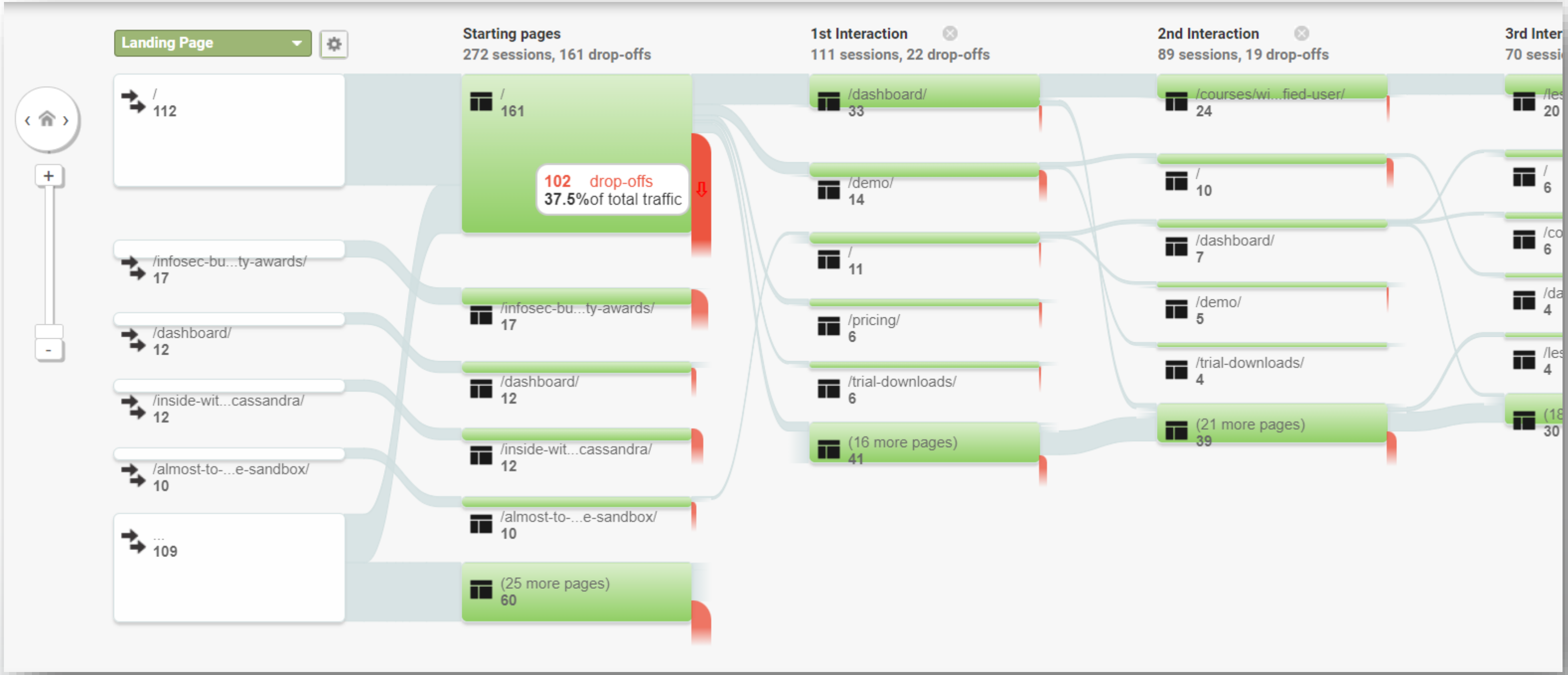
Key Tracking Metrics

- Agent ID (Preferred) or Cohort ID
- Content ID
- Action ID
- Action State
- Proposed Reward ID

ID	Cohort	CID	AID	AS	RID
1000	1	3	2	0	1
1002	2	3	2	1	1
1000	1	4	2	1	2
1001	2	3	2	1	1

https://test.com/page.php?agent_id=1001&content_id=3&action_id=2&state=fail&reward_id=4

Behavior Flow Analysis



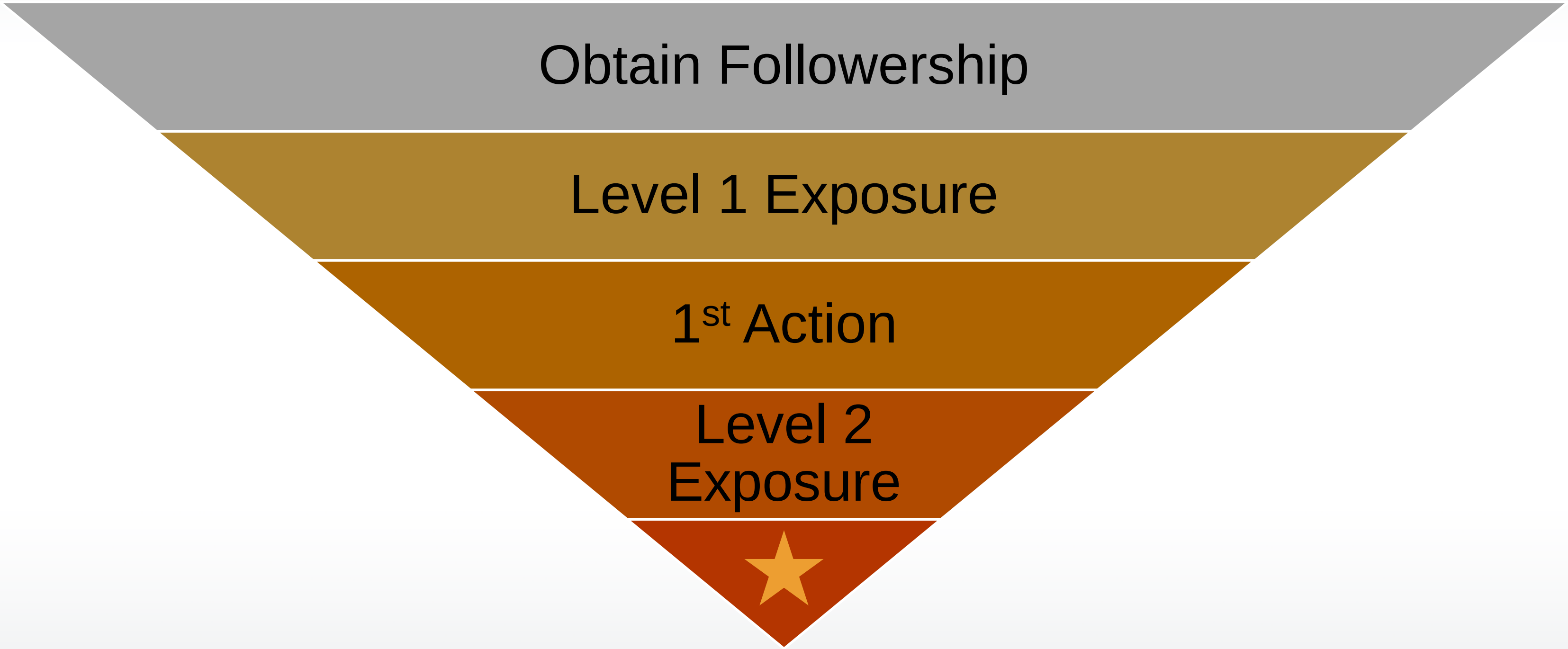
witfoo

Improving Cohort Definitions

- What characteristics are shared among funnel exits?
- Are those characteristics missing in continuations?
- Tracking interaction counts to conversion

ID	Age	Politics	Income	Role	Music
1003	57	Near Left	$120k	PM	EDM
1198	45	Far Left	$200k	Exec	EDM
2876	33	Near Right	$85k	Dev	EDM

Multi-stage Cultivation



Maintaining a Campaign

- Data Mining for Cohort Definition
- Large, diverse, effective content
- Ability to maintain audience (Groups, followers, mail lists, games)
- A/B testing of action delivery
- Reward types
- Decent Model (Naive Bayes)
- Growing Dataset for Model
- Detection Avoidance

Machine Learning Driven Social Engineering

Charles Herring
Co-Founder, CTO

me@CharlesHerring.com
https://CharlesHerring.com
@charlesherring