



Bot Detection

November 6, 2015

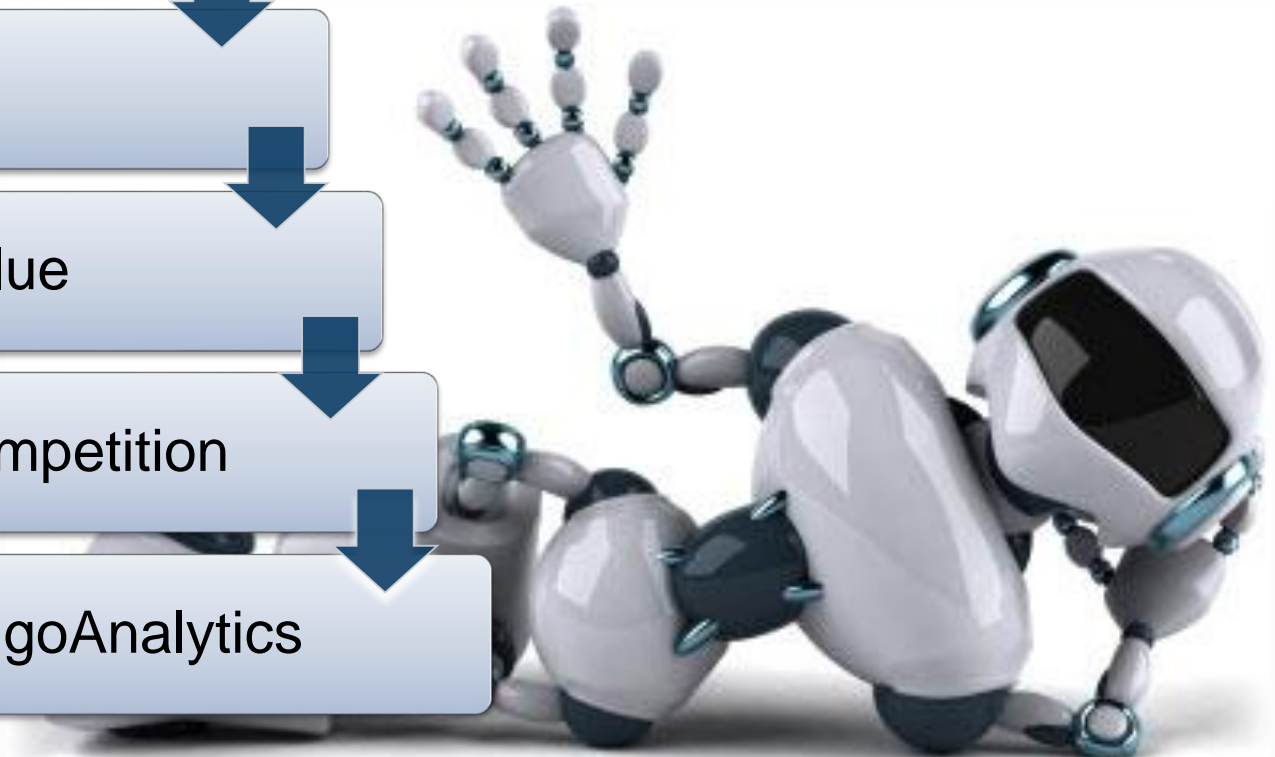
Introduction

Case study

Business value

Kaggle competition

About AlgoAnalytics



What's a bot?

- An application that runs automated repetitive tasks over the Internet at a much higher rate than a human
- Designed to be undetectable by mimicking human behavior

What's it used for?

- Fetching and analyzing files from web servers
- Sending rapid responses in online auction sites
- Manipulating reviews in e-commerce websites
- Committing fraud clicks on online advertisement
- Coordinating automated attacks on network computers

- **Click Fraud Defined:**

- Occurs in a pay-per-click scenario where a computer program imitates a web user by clicking on ads to generate a higher charge per click
- Bots used for invalid clicks to drive up advertisers' costs and eat up promotional budgets without generating any commensurate return

- Google and Yahoo! faced lawsuits due to fraudulent clicks:

- Yahoo paid \$4.5 million in legal bills for the plaintiffs and agreed to settle advertiser claims.
- Google settled a similar click fraud lawsuit with Lane's Gift and Collectibles for \$90 million.

Can these fraudulent clicks be detected in time?

Online Auction Fraud Examples



- In eBay, 65% of sessions are bot sessions.
- Bots are mainly used for
 - Manipulating feedback forms to get abundant positive reviews
 - Placing bids faster than humans, raising the bidding price or winning the auction
 - Hacking a trustworthy seller's account and starting a bogus auction

Bot Detection Models

- Dynamic!
 - Never rely on static signatures or fixed rules
- Machine Learning
 - Classify each session on features derived from historical behaviour data
 - Model self-tunes to evolve alongside the bot threats

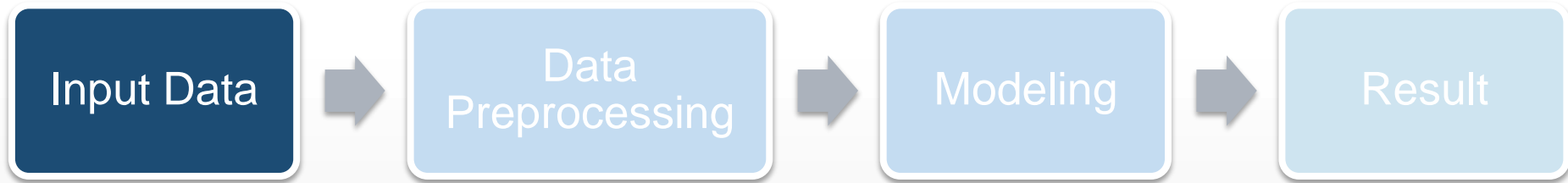


Goal

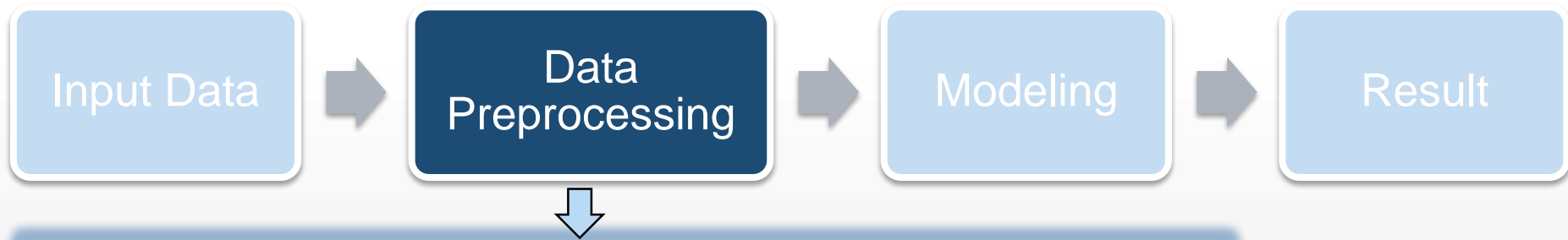
- To classify online auction bids as "bots" or "human".

Problem description

- Online auctions have become an increasingly important aspects of e-commerce.
- However, for consumers, the chances of landing a winning bid in online auctions has become more difficult due to the abundance of "bidding robots".
- By identifying bids that are placed by "bots", the site owners can remove these users and prevent unfair auction activity



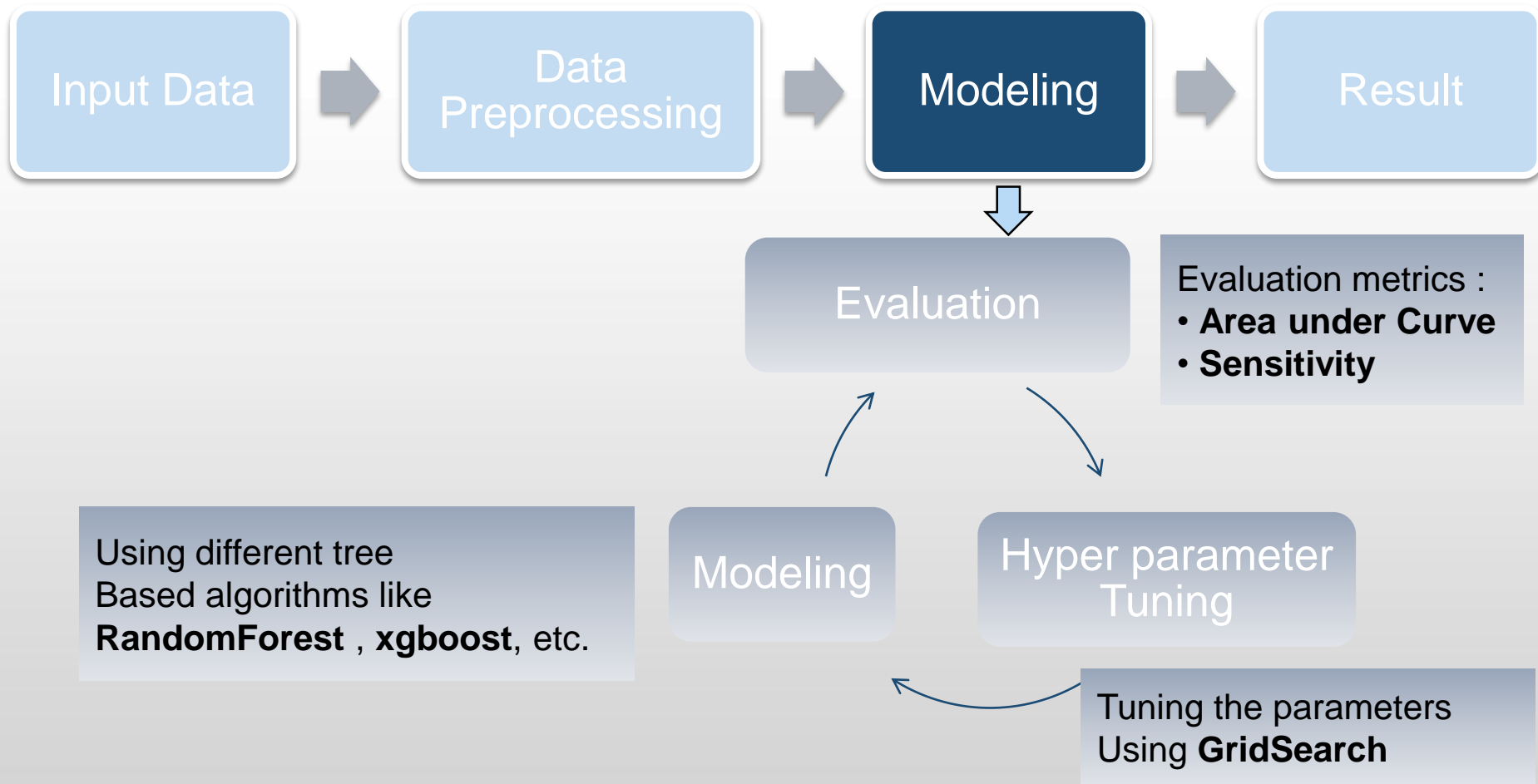
- **Train.csv** Bidder IDs with labels (Human or Bot)
- **Test.csv** Bidder IDs
- **Bids.csv** Information of each Bidder ID for each auction



For each bidder we create features such as:

- Variability in IP's used
- Variability in URL's
- Time between the current bid and the previous bid
- Total number of bids placed
- Time between a user's bid and the previous bid by another user in the same auction

Methodology :



Methodology



Completed • Jobs • 985 teams

Facebook Recruiting IV: Human or Robot?

Mon 27 Apr 2015 – Mon 8 Jun 2015 (4 months ago)

Dashboard

Private Leaderboard - Facebook Recruiting IV: Human or Robot?

This competition has completed. This leaderboard reflects the final standings.

See someone using multiple accounts?

[Let us know.](#)

#	Δrank	Team Name <small>‡ model uploaded</small>	Score <small>🔍</small>	Entries	Last Submission UTC (Best – Last Submission)
1	↑87	Life in a Glass House ‡	0.94254	3	Mon, 08 Jun 2015 10:20:49
2	↑4	small yellow duck ‡	0.94167	9	Mon, 08 Jun 2015 16:59:55 (-20.2h)
3	↑2	mechatroner ‡	0.94114	29	Sun, 07 Jun 2015 23:23:22 (-24.5h)
4	↓2	SY	0.94079	58	Fri, 05 Jun 2015 13:01:17
5	↑7	square7 ‡	0.93992	44	Mon, 08 Jun 2015 20:46:19 (-4.1h)
6	↑13	Mario Filho ‡	0.93964	31	Mon, 08 Jun 2015 14:39:40 (-4d)
-		Team_AlgoAnalytics	0.93961	-	Thu, 15 Oct 2015 10:05:33 <small>Post-Deadline</small>

Post-Deadline Entry

If you would have submitted this entry during the competition, you would have been around here on the leaderboard.

7	↑25	Artem R.	0.93939	36	Mon, 08 Jun 2015 09:52:35 (-3d)
8	↑78	ask788 ‡	0.93939	41	Mon, 08 Jun 2015 20:13:45 (-3.4h)

About AlgoAnalytics

▪ Areas of expertise

- Predictive analytics consultancy
- Algorithmic and quantitative trading in financial markets

▪ About us

- Founded by Dr. Aniruddha Pant in 2009
- Firm focused on the intersection of mathematics, computer science and other domains

▪ Team



Dr. Aniruddha Pant
CEO

Ph.D. (Berkeley, USA)

Control theory, Machine
Learning, Analytics,
Statistics



Rohit Walimbe
Sr. Quant Analyst

B.Tech, PGDM

Analytics, Finance



Mukta Pendharkar
Quant Analyst

MSc Financial Maths

Analytics, Finance

+ 6 Quantitative Analysts

with a solid foundation in mathematics and engineering