

Reverse Engineering how WAFs Like Cloudflare Identify Bots

Sociotechnical Security?

Security as an evolution from complication to complexity

But if “security” is to be applied to participatory technologies ... we run into some new challenges. These are complex systems that touch on a range of interconnected social dynamics, profit incentives, psychological variables, technology designs, data storage systems, automated and machine learning algorithms—and all the things that emerge from the messy interactions of those different things. This isn’t as simple as just finding bugs in code.

- Goerzen, Watkins, & Lim, Foci ‘19



Abuse is becoming the biggest game in town

TECH

How Sellers Trick Amazon to Boost Sales

The digital giant battles click farms, reviewers-for-hire and other scams as merchants try to outsmart its product-ranking system.

ELECTION 2020

Fake Twitter Accounts Posing as News Organizations Prematurely Declare Election Victories

Similarities in fake accounts and tweets point to coordinated effort to spread online disinformation about presidential election

The New York Times

Facebook Identifies an Active Political Influence Campaign Using Fake Accounts

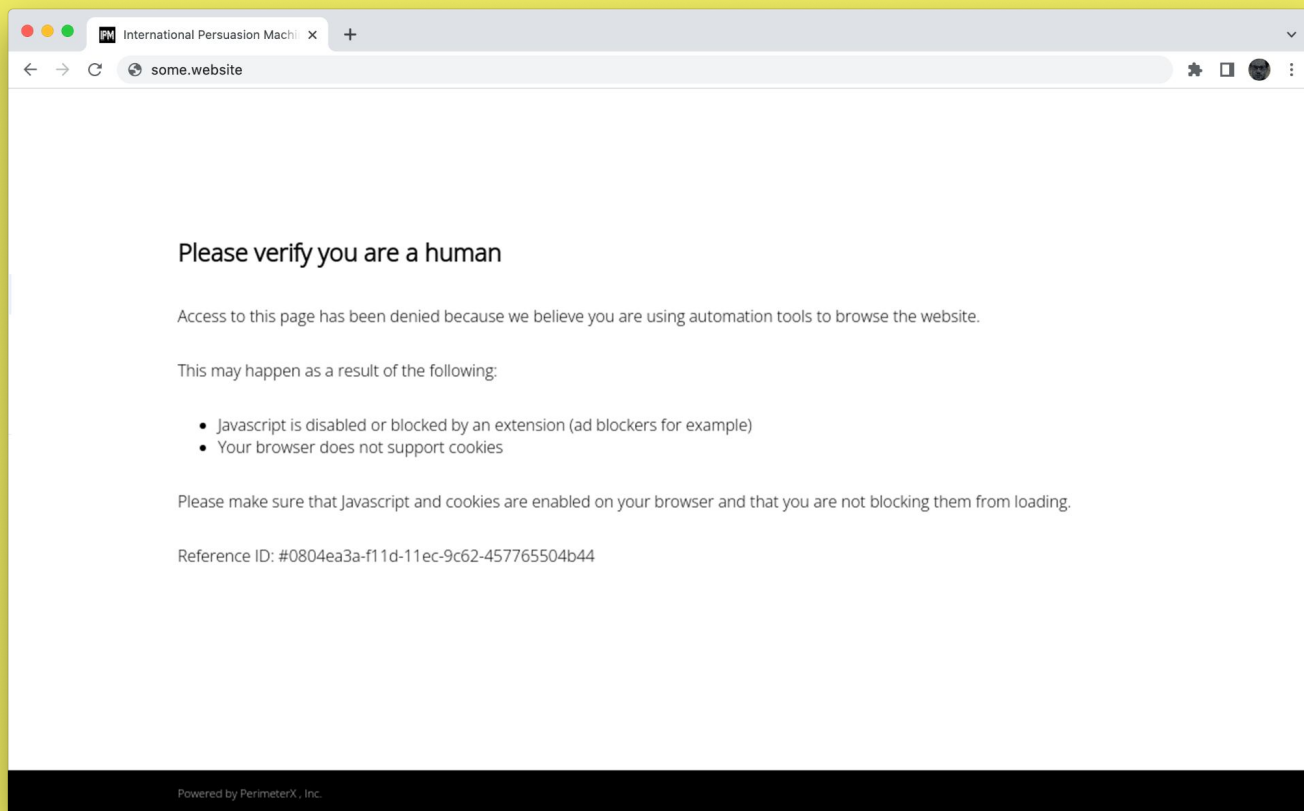
TECH / ARTIFICIAL INTELLIGENCE

AI trained on Yelp data writes fake restaurant reviews 'indistinguishable' from real deal

'Omgggg! Very flavorful!! It was so delicious that I didn't spell it!!'

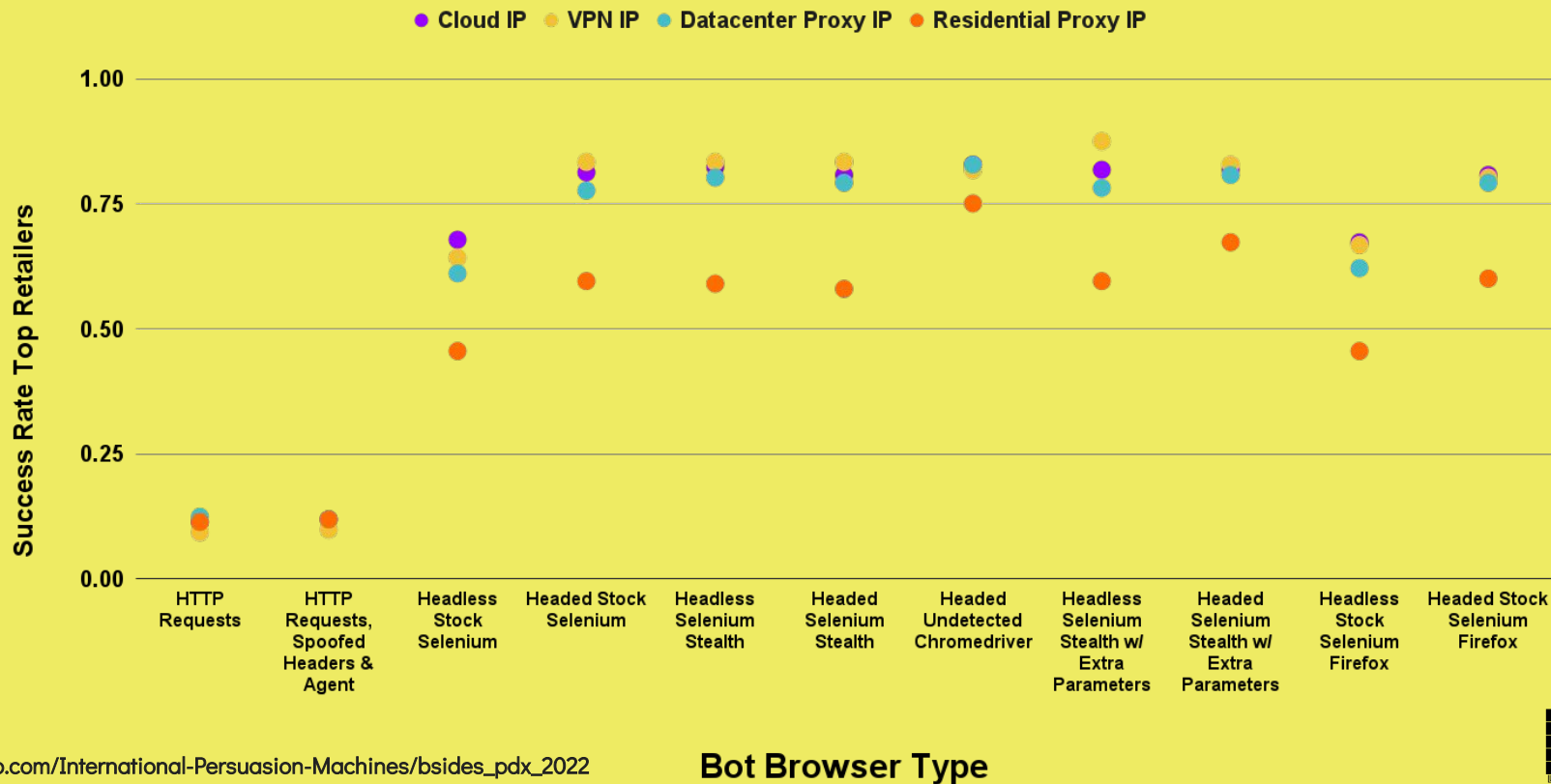
By James Vincent | Aug 31, 2017, 5:39am EDT

WAFs as Bot Security Theater















































WAFs as Bot Security Theater

Browser Attack Performance, Browser x Network Interfaces



What makes WAFs Tick?

	Cloud IP	VPN IP	Datacenter Proxy IP	Residential Proxy IP
Plain HTTP Requests				
HTTP Requests, Spoofed Params				
Headless Stock Selenium				
Headed Stock Selenium				
Headless Selenium Stealth				
Headed Selenium Stealth				
Headed Undetected Chromedriver				
Headless Stealth w/ Extra Params				
Headed Stealth w/ Extra Params				
Headless Stock Selenium Firefox				
Headed Stock Selenium Firefox				



Obfuscates IP



Obfuscates Browser



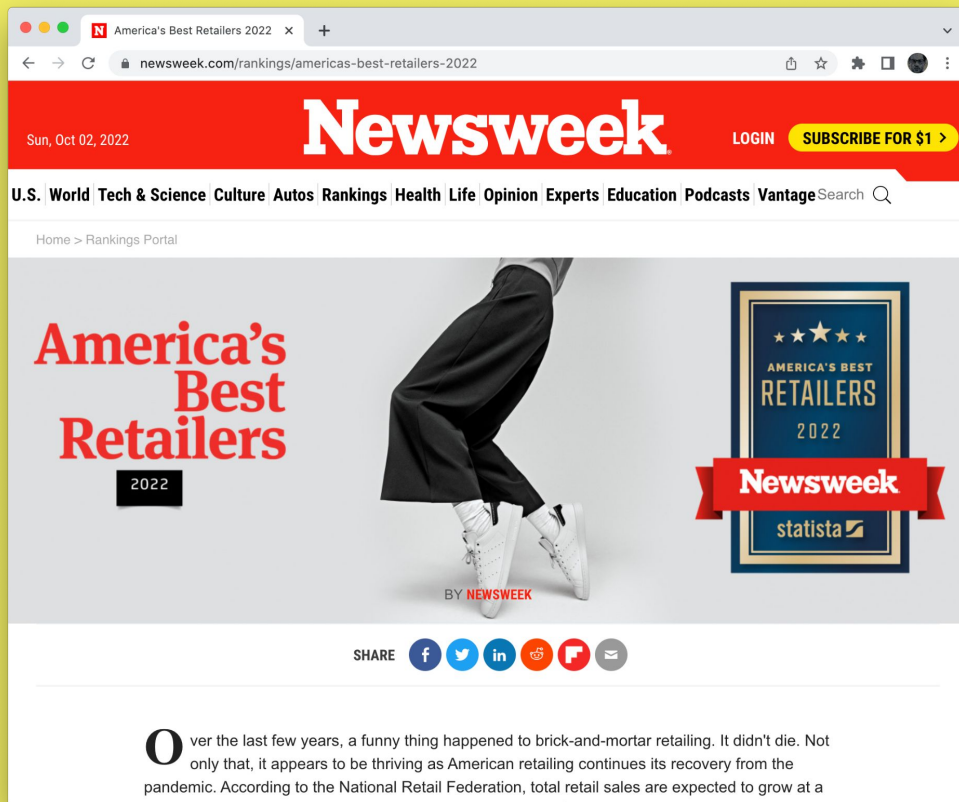
Virtual Browser



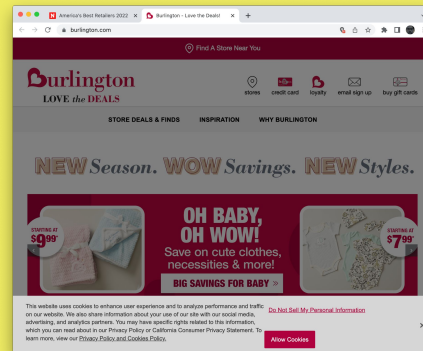
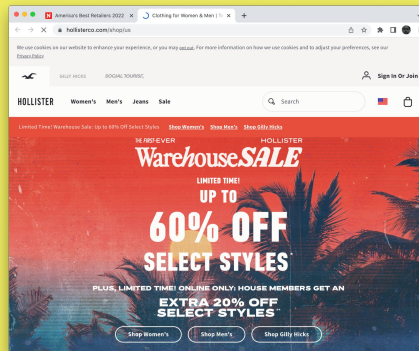
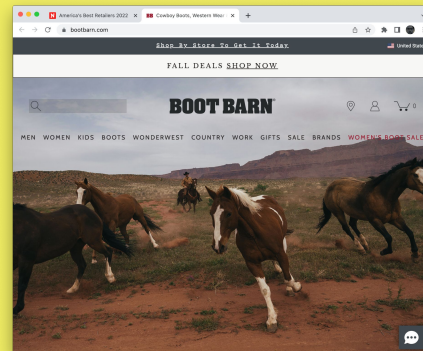
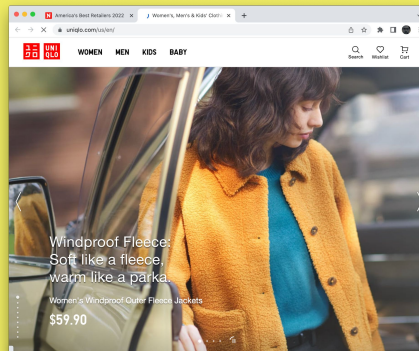
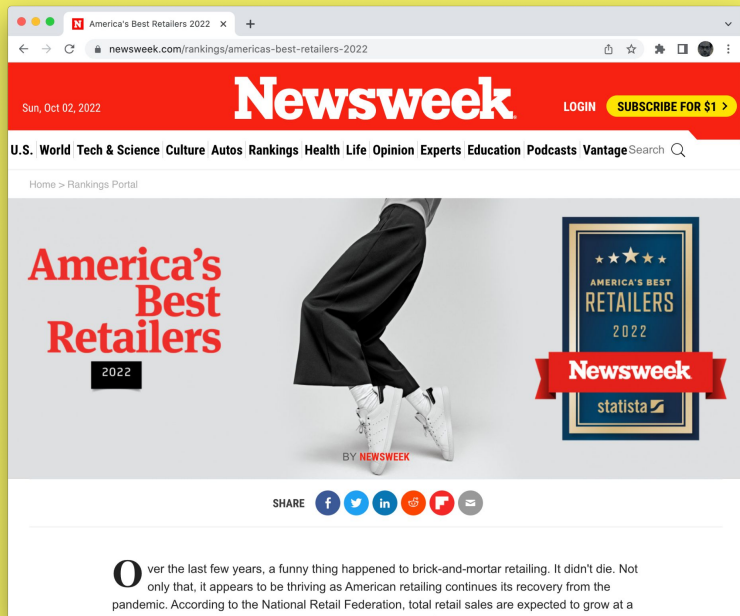
Virtual Display



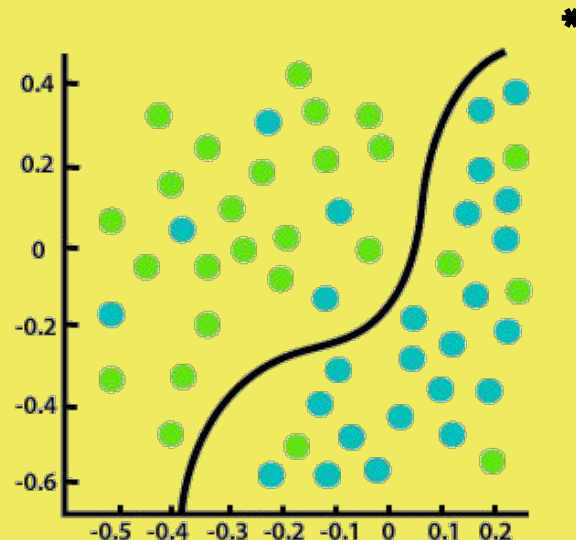
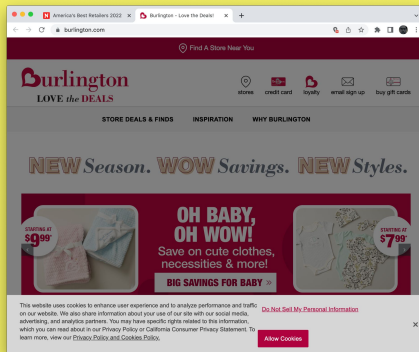
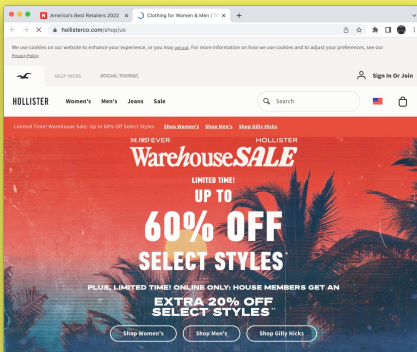
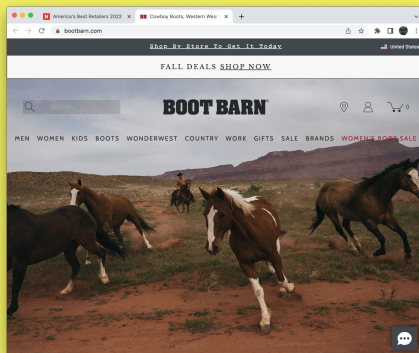
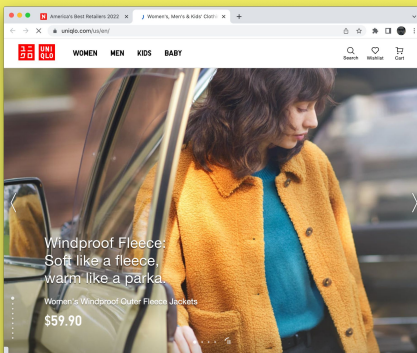
Reverse Engineering WAFs: Dataset



Reverse Engineering WAFs: Dataset

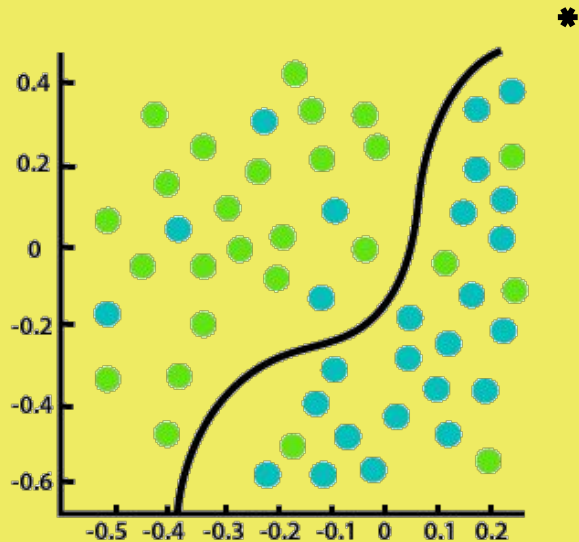


Reverse Engineering WAFs: Labeling



* 96% Accuracy, 316 features, 1,608 labeled cases

Reverse Engineering WAFs: Feature Encoding

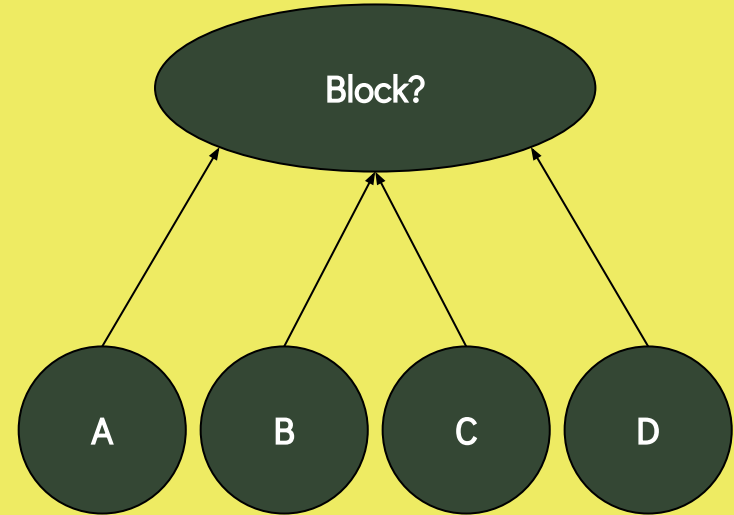


Blocked?	A?	B?	C?	D?
1	0	1	1	1
0	1	0	0	1
1	1	0	1	0

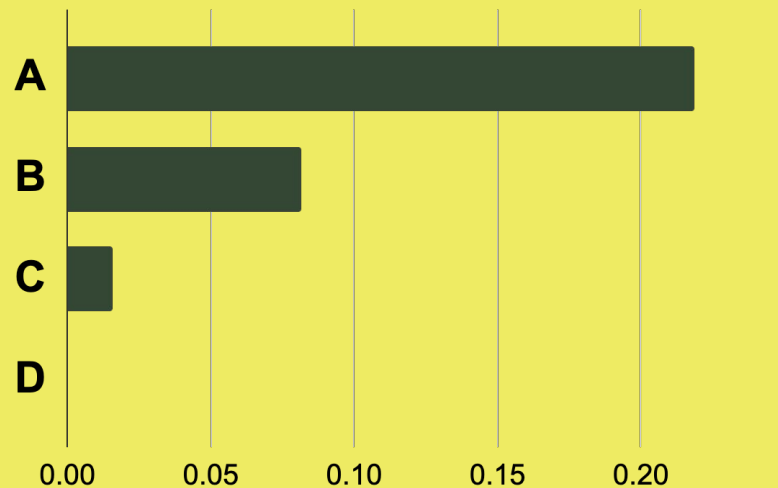
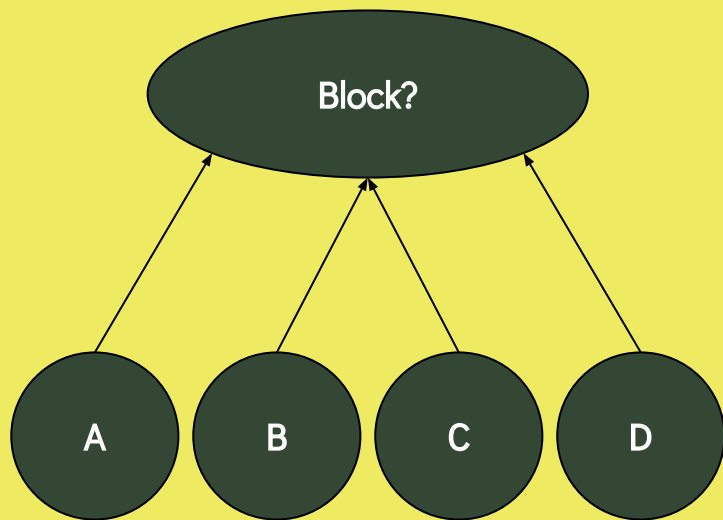
* 96% Accuracy, 316 features, 1,608 labeled cases

Reverse Engineering WAFs: Dominance Analysis

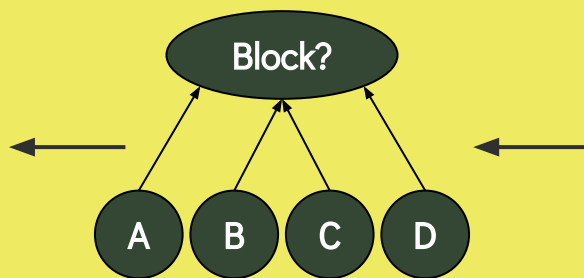
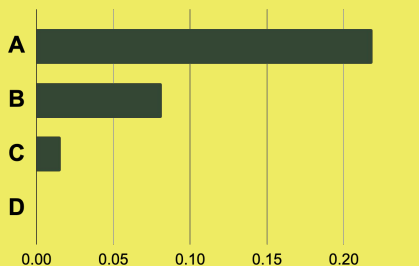
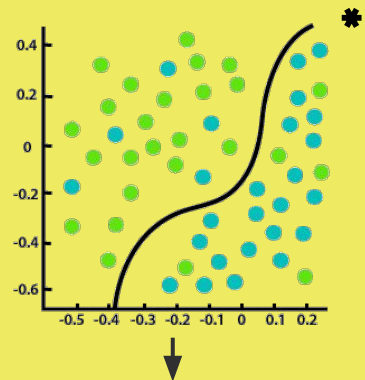
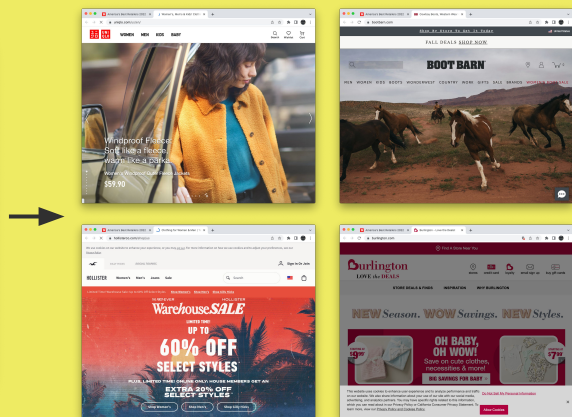
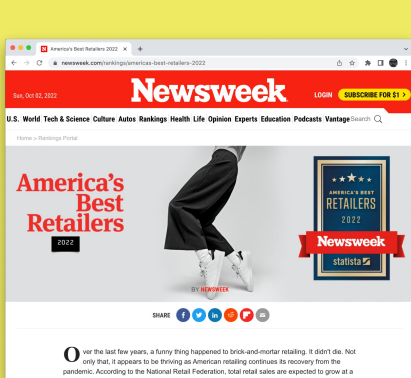
Blocked?	A?	B?	C?	D?
1	0	1	1	1
0	1	0	0	1
1	1	0	1	0



Reverse Engineering WAFs: Results



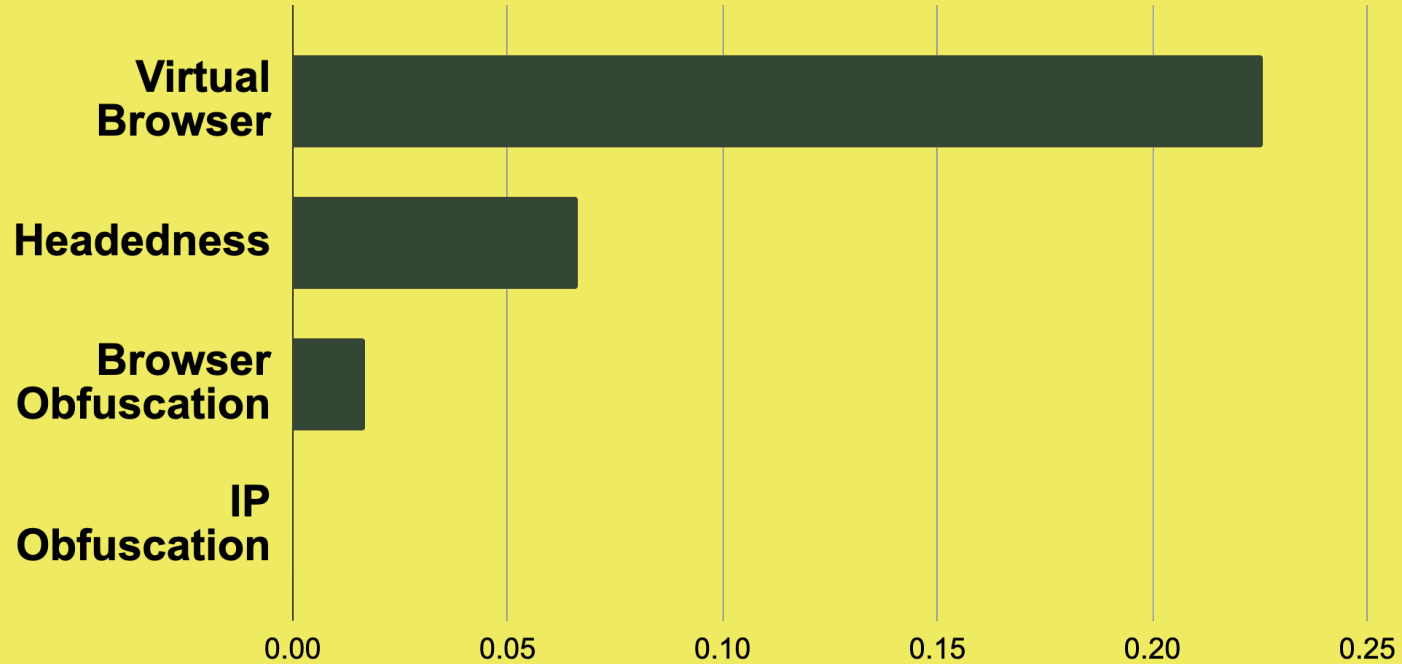
Reverse Engineering WAFs: Full Model



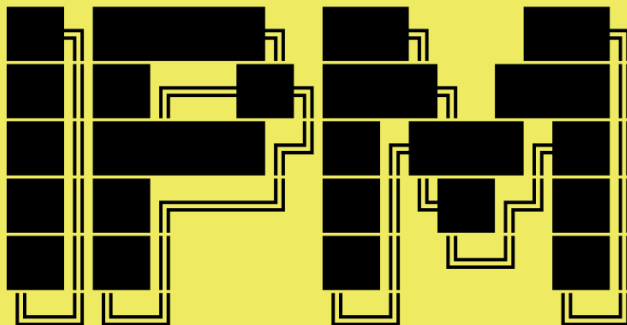
Blocked?	A?	B?	C?	D?
1	0	1	1	1
0	1	0	0	1
1	1	0	1	0

* 96% Accuracy, 316 features, 1,608 labeled cases

WAF Dominance Analysis Relative Weighting



Relative Factor Importance ($R^2=0.31$)



@intl_persuasion

ipm-corporation.com