



Technical Showcase:

How Netacea Develops Cutting-Edge Bot Detection Techniques

THE EVER-EVOLVING BOT PROBLEM

On every website there are bots making automated requests. Bot traffic is now so prevalent that the internet as we know it today could not function without bots, for better or worse.

Although some bots are beneficial, malicious bots are using a whole host of methods to make a profit at the expense of websites and users. From scraping to scalping, account takeover and creation to carding and credential stuffing, the threats posed by malicious bot traffic are growing more dangerous and harder to defend against.

These attacks are distinct from traditional hacks that exploit technical weaknesses in systems. Bots automate legitimate activity on websites to exploit business logic weaknesses, making conventional security measures ineffective against these threats.

To compound the problem, modern bots are sophisticated enough to bypass client-side detection techniques used by many bot defence systems, such as device fingerprinting and user interaction tracking.

Netacea's Intent Pathways technology saves customer £3 million

INTENT PATHWAYS IN ACTION: Stopping odds scraping and arbitrage betting on a gaming site

A global gaming and betting organisation had a problem with bots continually scraping their website for odds. Arbitrage bettors used this information to place bets on every outcome of specific events, guaranteeing a profit. Not only does this cheat the system, but these bots represent a significant proportion of overall traffic, so serving them adds significant costs to the website's infrastructure.

Netacea needed to consistently identify this scraping behaviour in real-time using the huge amounts of data we collect, which was unlabelled and hard to classify consistently across websites and threat types.

Our data science team observed that only by looking at the request-by-request level of website interactions does the individual behaviour of visitors become apparent. This makes it difficult for machines to identify scraping behaviour, especially at scale and in real-time.

LEVERAGING DATA SCIENCE TO BEAT THE BOTS

To innovate new solutions whilst providing value to customers, Netacea's data science team follows the 3X methodology (explore, expand, extract) pioneered by Agile evangelist Kent Beck. This allows us to rapidly test ideas, scale them and get buy-in from customers before extracting a finished product.

By validating our ideas with customers and collaborating with them when building new products, we can innovate efficiently and provide real value to our customers.

The 3X Methodology



Explore new ideas



Expand on findings with customers



Extract into new products and features

FIRST X: EXPLORE

The first phase of the 3X approach, “explore”, is designed to quickly experiment with lots of different ideas to find one that sticks. In this case, the winning idea was to compare our problem with that of natural language processing.

In natural language processing, a neural network is trained to predict the probability of words (or “tokens”) in the English language appearing together. After embedding the relationships between the 170,000 words in our vocabulary, sentences can be analysed computationally to evaluate metrics, for example the overall sentiment from a dataset of tweets.

For an eCommerce website, the “vocabulary” is made up on average of 100,000+ pages, split into individual paths, and sequenced as a session or user journey between those requests. We wanted to know whether we could train the embeddings of paths, which allow us to analyse the intent of users based on the sequence of requests they were making, to decide in real-time whether visitors were scrapers or genuine customers.

We called this new technology “Intent Pathways”, because it allows us to identify and categorise the digital footprints that users leave as they make requests and navigate between paths on the website.

Learn how Netacea applied natural language processing methodologies to accurately detect bot traffic in real-time in an on-demand webinar with Matt Jackson, Head of Data Science.



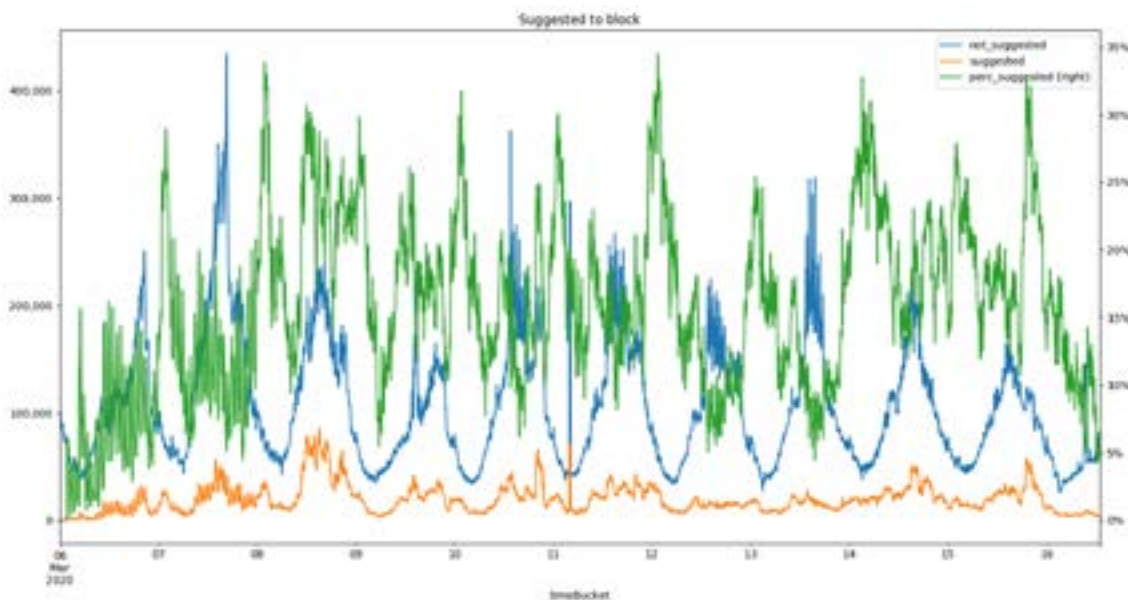
The results of our initial tests were very encouraging. A 3D representation of the output revealed clear clustering of related paths on our target website, meaning the neural network had learned to contextualise this information.

SECOND X: EXPAND

After the successful experiment we moved onto the second “X”, expand, to see if this could provide value to our customer. We collaborated with our gaming and betting customer to identify “footprints” likely to signify scraper activity and applied this to our live monitoring of their website, alongside our existing batch classification model.

After running the test for several days, we found that most of our recommendations for blocking scraper bots came in real-time from our new “Intent Pathways” product. This new technology revealed that a large percentage of the total traffic was scraper bots.

By blocking these requests in real-time, we not only reduced unwanted bets attempted by bots by 85%, but also reduced the overall number of website requests by 40%. All this added up to overall savings of £3 million for the customer across infrastructure, fraud losses and staffing.



A significant percentage of traffic was identified as bots and blocked for the client.

RECOMMENDATIONS

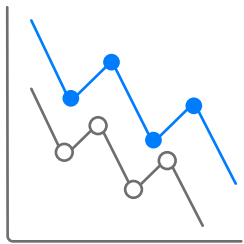


The vast majority of blocking recommendations were made by the new Intent Pathways technology.

THIRD X: EXTRACT

Using the 3X approach, we explored a challenging problem, developed a unique solution and provided significant value to a customer whilst collaborating directly with them in its development. This also built trust with our customer and extracted a useful feature we are already using with other customers.

Intent Pathways (patent pending) is now one of Netacea's core technologies and is uniquely embedded in our bot detection and mitigation product.



40%

reduction in total
website requests



£3 million

Overall savings across
infrastructure, fraud
losses and staffing



85%

reduction in unwanted
bets placed by bots

THE SOLUTION TO YOUR BOT PROBLEMS

At Netacea we are committed to providing our customers with the most up-to-date innovations in mitigating harmful bot traffic. We do this by linking our proactive data science team with the development of new product features, working closely with customers to provide real value with unique technology.

This mindset, paired with our server-side approach and innovative Intent Analytics™ technology, allows us to seamlessly integrate with your business and deliver accurate, intelligent and effective bot mitigation.

To find out more about Netacea's unique approach to stopping sophisticated bot threats, visit www.netacea.com/why-netacea or talk to our team today at hello@netacea.com.

