

Electric Bike for the Brain

Let's poke fun at AI

Awfully performing V0-generated React

<https://fontes-locais.vercel.app/>

AI sucks!

Stages of grief

denial, anger... acceptance

Stages of AI grief

- Oh S#!^, it will take my job (panic)
- Nah, never, it's not good enough (denial)
- Maybe it could one day take my job
- You know what, if it could, it should
- Please, take my job!

AI is the best!

About me

- **Stoyan Stefanov**
- **Web performance enthusiast, since 2007?**
- **Yahoo, Facebook, WebPageTest, Etsy**

Web perf work

- **Investigate**
- **Research**
- **Make test pages**
- **Prototype**
- **Ship code to production**

Steve Jobs on computers

Bicycles for the mind



AI: an *electric* bike for the mind

**But... AI-gen'd production code
in my web perf day-to-day?**



**NO BIKES
BEYOND THIS
POINT**

MCP: Model Context Protocol

Chrome devtools

Your IDE now has access to the browser

Users > ssstefanov > .cursor > {} mcp.json > ...

```
1  < {  
2    < "mcpServers": {  
3      <   "chrome-devtools": {  
4        <     "command": "npx",  
5        <     "args": ["-y", "chrome-devtools-mcp@latest"]  
6      }  
7    }  
8  }
```

“Open the browser, navigate to etsy.com, run a performance trace and give me top 3 areas for improvement”

“Do it again, but throttle Network and CPU”

MCP

- It lies
- You must keep asking questions

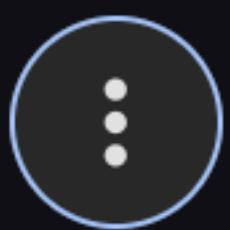
Keep asking questions

- In general when interacting with AI tools
- AI happily generates new code
- It needs YOU to nudge towards good engineering practices
- Including performance

Fixing the VO code

How about some DIY?

Gemini



PRO



what did Steve jobs said exactly about the computer being
the bycicle of themind



Normally I can help with things like this, but I don't seem to have access to that content. You can try again or ask me for something else.



"Rich and absorbing...a revelation." —Farhad Manjoo

BRIAN
MERCHANT

BESTSELLING AUTHOR OF *THE ONE DEVICE*

BLOOD
IN THE
MACHINE



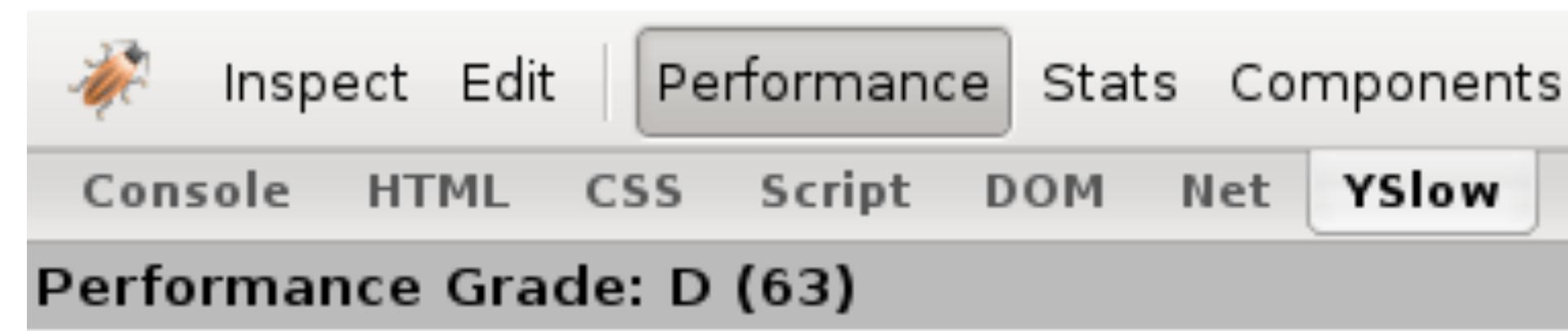
THE ORIGINS OF THE REBELLION
AGAINST BIG TECH

Let's develop some AI

ySlow
ySlow, PageSpeed, Lighthouse

ySlow “rules”

The Original 13



The screenshot shows the YSlow extension interface within a browser's developer tools. The top navigation bar includes tabs for Inspect, Edit, Performance (which is selected), Stats, Components, Console, HTML, CSS, Script, DOM, Net, and YSlow. A prominent message at the top states "Performance Grade: D (63)". Below this, a numbered list of 13 performance optimization steps is displayed, each with a grade (B, F, C, A, or n/a) and a description:

- B 1. Make fewer HTTP requests ▶
- F 2. Use a CDN ▶
- F 3. Add an Expires header ▶
- C 4. Gzip components ▶
- A 5. Put CSS at the top
- A 6. Put JS at the bottom
- A 7. Avoid CSS expressions
- n/a 8. Make JS and CSS external ▶
- B 9. Reduce DNS lookups ▶
- C 10. Minify JS ▶
- A 11. Avoid redirects
- B 12. Remove duplicate scripts ▶
- A 13. Configure ETags

aiSlow

Eh, I slow?

Naïve take

Take a bunch of HARs and spit out wisdom, please!

Easy there

Raw messy JSON data is no good, Excel-like needed

Target vs Features

Crawl the web and collect!

Nah, HTTPArchive instead

Crawls million of pages monthly

Data is in BigQuery

Get some data

httparchive / Datasets / crawl / Tables / pages

pages [Query](#) [Open in ▾](#) [Share ▾](#) [Copy](#) [Snapshot](#) [Delete](#)

Schema	Details	Preview	Insights	Lineage	Data Profile
<input type="checkbox"/>	date	DATE	REQUIRED	YYYY-MM-DD format of the HTT...	▼
<input type="checkbox"/>	client	STRING	REQUIRED	Test environment: desktop or m...	▼
<input type="checkbox"/>	page	STRING	REQUIRED	The URL of the page being tested	
<input type="checkbox"/>	is_root_page	BOOLEAN	REQUIRED	Whether the page is the root of ...	▼
<input type="checkbox"/>	root_page	STRING	REQUIRED	The URL of the root page being ...	▼
<input type="checkbox"/>	rank	INTEGER	NULLABLE	Site popularity rank, from CrUX	
<input type="checkbox"/>	wptid	STRING	NULLABLE	ID of the WebPageTest results	
<input type="checkbox"/>	payload	JSON	NULLABLE	JSON-encoded WebPageTest re...	▼
<input type="checkbox"/>	summary	JSON	NULLABLE	JSON-encoded summarization ...	▼
<input type="checkbox"/>	▶ custom_metrics	RECORD	NULLABLE	Custom metrics from WebPageTest	
<input type="checkbox"/>	lighthouse	JSON	NULLABLE	JSON-encoded Lighthouse report	
<input type="checkbox"/>	▶ features	RECORD	REPEATED	Blink features detected at runti...	▼
<input type="checkbox"/>	▶ technologies	RECORD	REPEATED	Technologies detected at runti...	▼

Naïve take

Pick 10K random sites and learn!

Easy there

Data is all over the place

Better: take a bunch of **slow** pages and a bunch of **fast** pages and learn

Filter outliers

- **too fast (error pages, forbidden, N/A in your country)**
- **too slow (temp issues, timeouts)**

Target & Features

Target: SpeedIndex

Features: 60-ish pieces of data

Learning

Regression (predicting a number) vs. **Categorization** (cat/dog)

VS LLM

VS Deep Learning

LightGBM: Light Gradient Boosting Machine

- A gradient boosting framework developed by Microsoft that uses tree-based learning algorithms
- **LGBMRegressor** is a specific model class (vs Categorization)

LightGBM

- Take structured CSV data with web performance metrics (bytes, requests, etc.)
- Train a regression model to predict SpeedIndex
- Use decision trees to find patterns in the data

... as opposed to Deep Learning

LightGBM (Gradient Boosting)

- **Architecture:** Ensemble of decision trees
- **How it works:** Builds trees sequentially, each correcting errors from previous trees
- **Best for:** Structured/tabular data (CSVs with rows and columns)
- **Training:** Fast (seconds to minutes)
- **Interpretability:** High - can see feature importance, SHAP values
- **Data needs:** Works well with small to medium datasets (thousands of rows)

Deep Learning

- **Architecture:** Neural networks with multiple layers
- **How it works:** Layers of interconnected neurons that learn patterns
- **Best for:** Unstructured data (images, text, audio, video)
- **Training:** Slow (hours to days, needs GPUs)
- **Interpretability:** Low - “black box” behavior
- **Data needs:** Usually requires large datasets (millions of examples)

Let's go!

`train_model.py`

Input-magic-output

Input

- Desktop-only
- Target: SpeedIndex
- Features: 60-ish pieces of data
- 20K+ fast pages (SI 500 - 2500ms)
- 20K+ slow pages (SI 4500 - 20000ms)

Magic

- **80/20 train/test split (`test_size=0.2`)**
- **Categorical Features**
- **DataFrame**

Magic

- **random_state=42 # reproducible results**
- **n_estimators=500 # Maximum number of boosting rounds**
- **learning_rate=0.05 # More careful, gradual learning**
 - Prediction = Tree1 + 0.05 * Tree2 + 0.05 * Tree3 + ...
- **stopping_rounds=10 # So we can exit before 500 rounds**

Output

- The Trained Model, aka The Brain
- `aislow_desktop.pkl`
- serialized model file, readable

Understanding the output

- **Training took 1.20 seconds.**
- **Best iteration: 111**
- **Model R-squared score on test data: 0.5754**
- **Feature importance**

Feature importance

- **Gain >> Splits > Correlation**
- **Gain %: Best indicator of what the model actually relies on**
- **Splits: Shows usage frequency (but not impact)**
- **Correlation: Shows linear trends (but model sees more than just linear)**

**Will it replicate ySlow's original
13?**

Feature	Splits	Gain %	Correlation	Impact
Total Page Size (Bytes)	201	44.84	0.355	↑↑ SLOWS page
First Paint	354	23.92	0.511	↑↑ SLOWS page
JavaScript Requests (Count)	360	4.28	0.367	↑↑ SLOWS page
Time To First Byte (ms)	162	3.51	0.339	↑↑ SLOWS page
JavaScript Size (Bytes)	117	2.27	0.486	↑↑ SLOWS page
Async Scripts	222	2.27	0.249	↑ slows page
Long Tasks Count	154	1.94	0.368	↑↑ SLOWS page
Inline Scripts	185	1.92	0.171	↑ slows page
Image Size (Bytes)	105	1.28	0.288	↑ slows page
Total Blocking Time	101	1.14	0.348	↑↑ SLOWS page
Max Requests Per Domain (Count)	51	0.95	0.327	↑↑ SLOWS page
CSS Requests (Count)	102	0.87	0.246	↑ slows page
Total Scripts	32	0.58	0.307	↑↑ SLOWS page

aiSlow

“A web-perf consultant in a box”

Given a page and the model, tell me what to improve

Also needs the training data (or a representative sample)

SHAP

SHapley Additive exPlanations

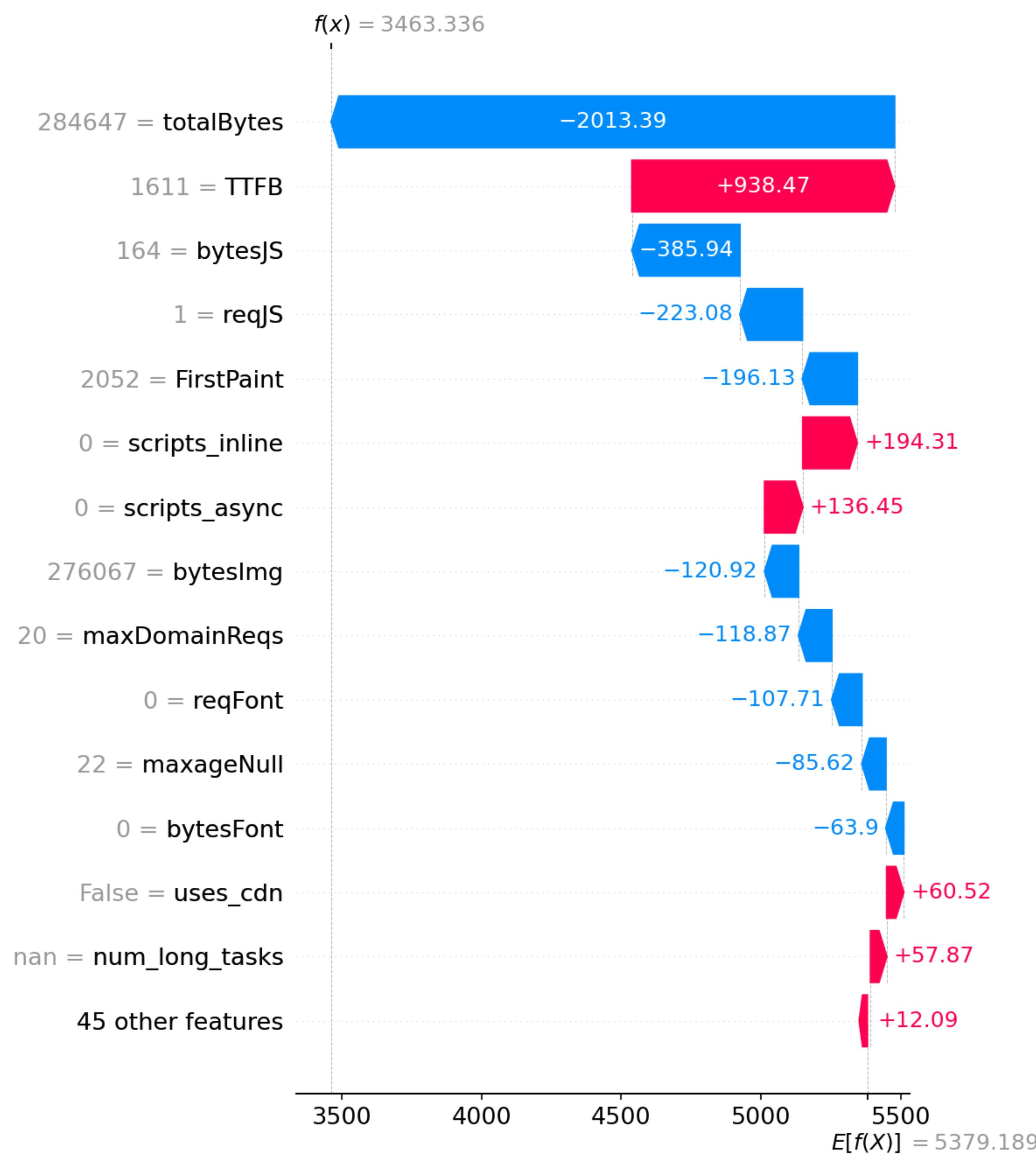
Technique for explaining machine learning model predictions

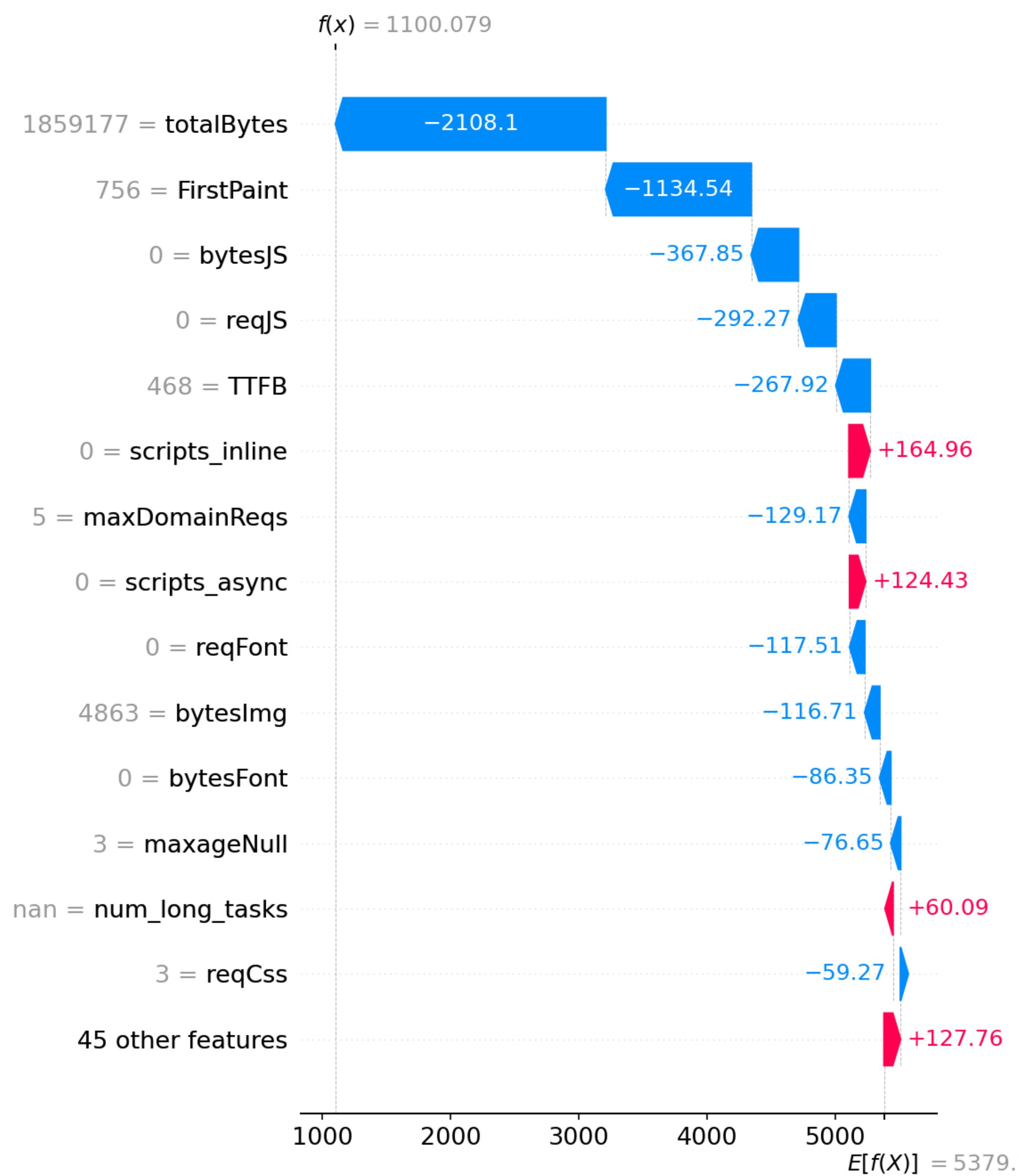
Game theory: how much each player contributed in a game

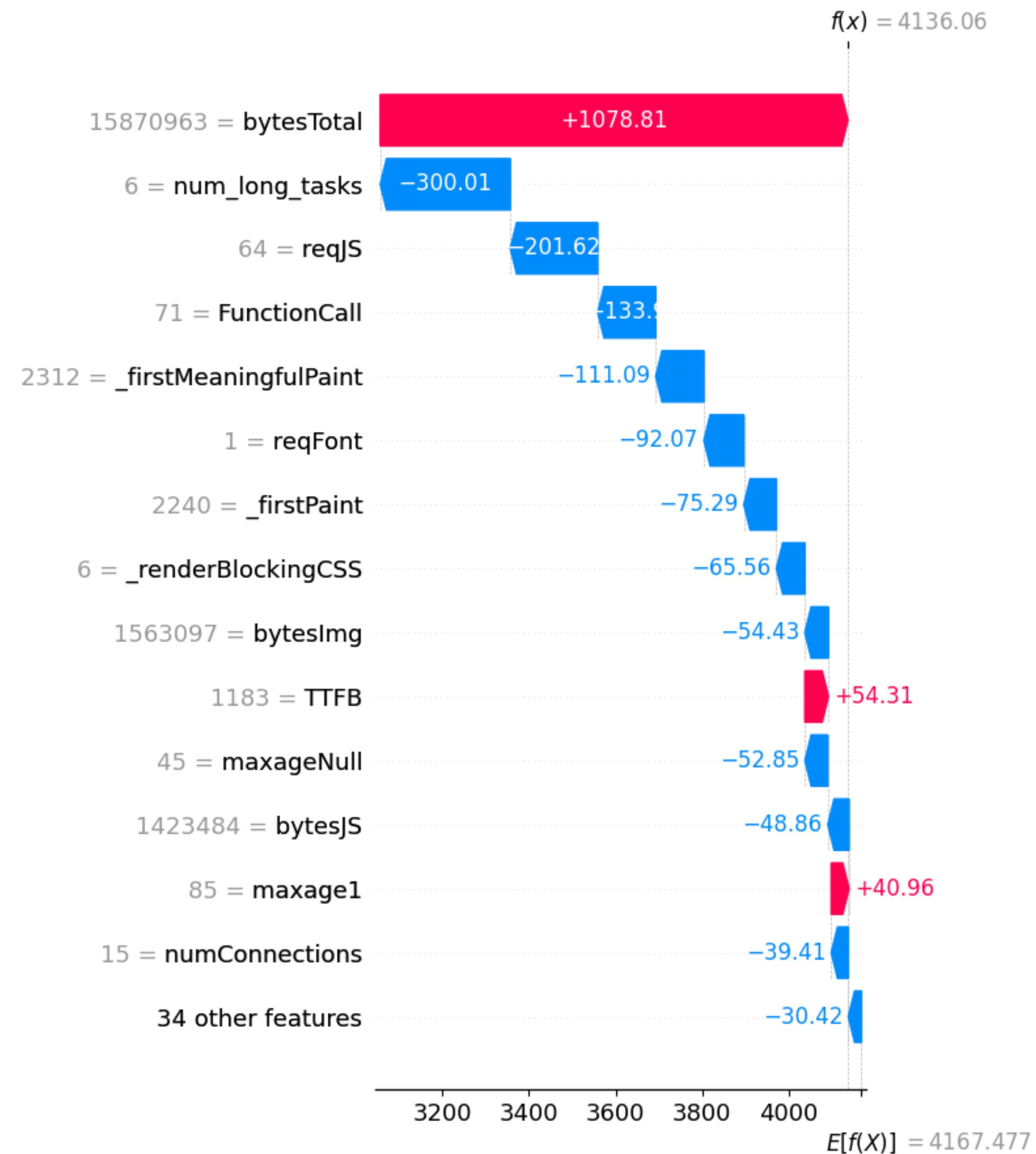
SHAP waterfalls

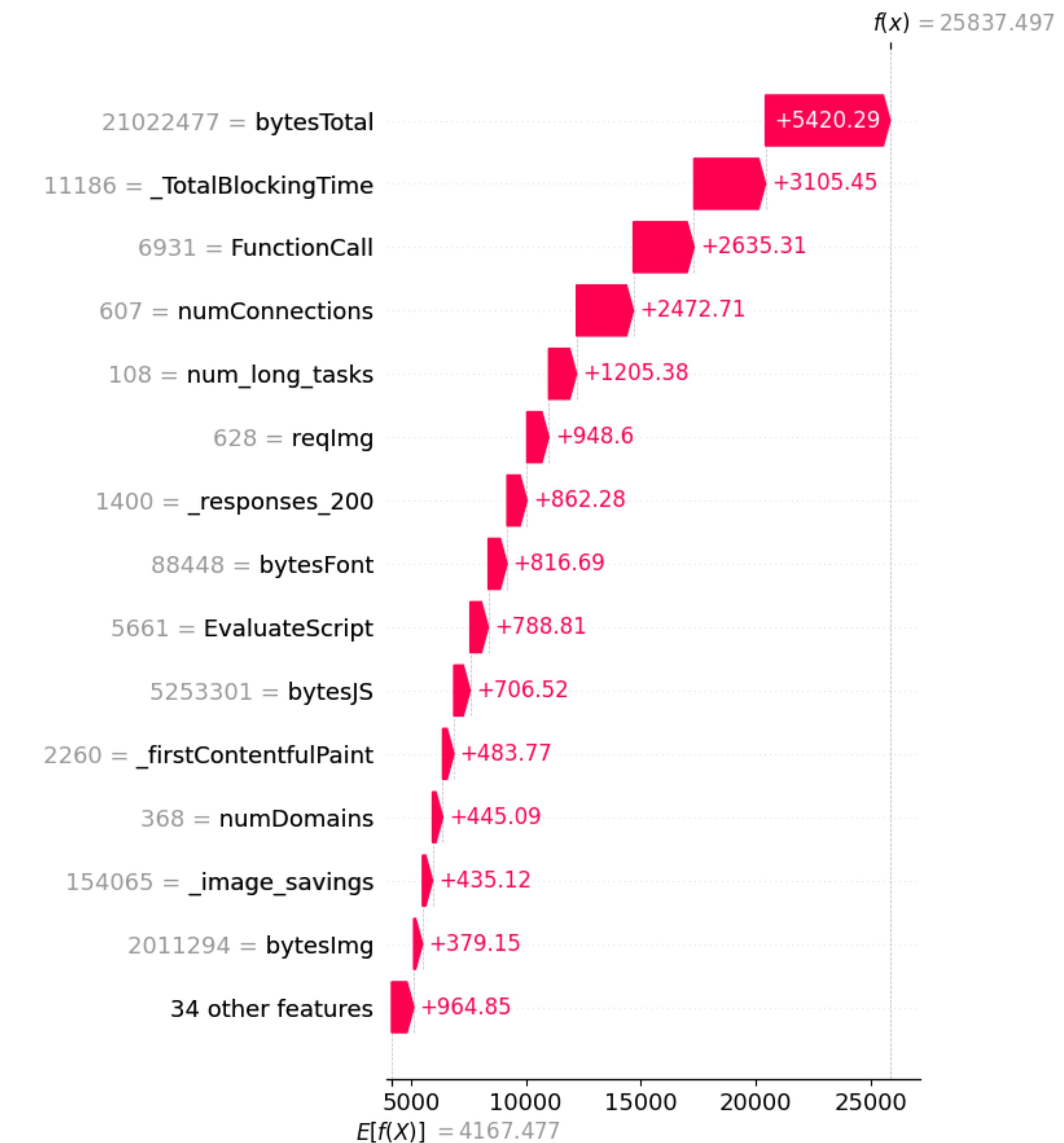
Start with the average in the data set

For each feature, calculate its impact, + or -









aiSlow

Show the top 3 + and top 3 - features

50ms+ only to reduce noise

What-if scenario

Take the top #1 worst offender

Make its value be the 25 percentile

Recalculate a new prediction

Examples

Will it take my job?

Eyes on the price: faster web

Plug

<https://calendar.perfplanet.com>



Web Performance Calendar

The speed geek's favorite time of year

2024

31st

Dec 2024

**Breaking Up with Long Tasks
or: how I learned to group
loops and wield the yield**

by [Rick Viscomi](#)



Everything, On the Main Thread, All at Once Arrays are in every web developer's toolbox, and there are a dozen ways to iterate over them. Choose

Thank you!

<https://bsky.app/profile/stoyan.me>