

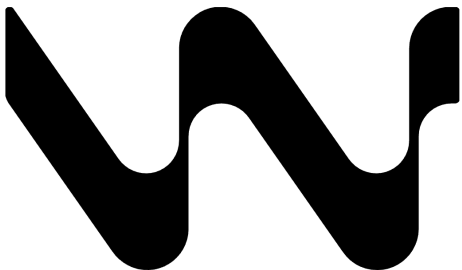
Building WebOpt environment

How LLM can optimize web app?

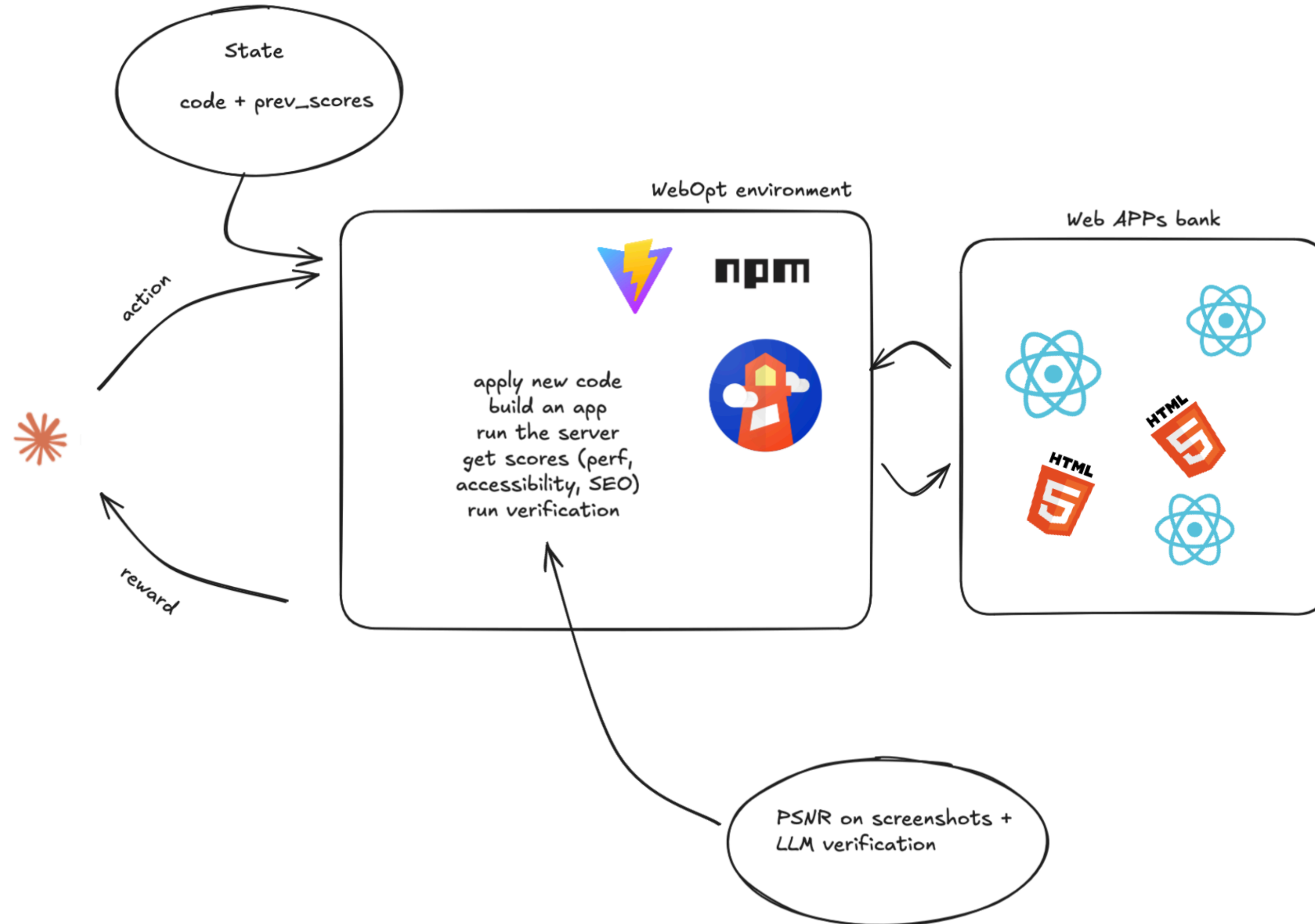
- Just by editing the code?
- ...But it's not enough

We've built the better environment
for optimizing the web :)

WebOpt env

- Based on openEnv standard
- Checks state of the art metrics for the web app
- Scores the whole pipeline, not just a code
- 30% vibecoded with 🕶️ 

How it works?



Projects bank

- The environment contains a bank of web application
- You can add new applications to the bank to make your agents scale

Rewards

- Average lighthouse metrics for perf, seo and accessibility
- Visual similarity evaluation (verification)
- Functional evaluation with Claude: (verification)
 - Create a specification of the original web app based on the screenshot
 - Score the new version of the app against the specification

Why it's useful?

- Works e2e. Modifications of the build configs will be scored as well
- Mixed reward function allows to adjust the reward based on web app structure, optimize for SEO or performance

Claude baseline

```
Initial Reward: 0
Modifying file: /tmp/current\_project/src/pages/Home.jsx
Reward: 0.6112568039520013
Step 1 completed.
=====

Modifying file: /tmp/current\_project/src/pages/Home.css
Reward: 0.0
Step 2 completed.
=====

Modifying file: /tmp/current\_project/src/pages/Home.jsx
Reward: -0.6276425293325595
Step 3 completed.
=====

Modifying file: /tmp/current\_project/src/pages/Home.css
Reward: 0.0
Step 4 completed.
=====

Modifying file: /tmp/current\_project/src/main.jsx
Reward: 0.2902594403503856
Step 5 completed.
=====
```

Check this out

<https://github.com/alex-kharlamov/WebOptEnv>

The image shows a web browser window with two main panels. The left panel, titled 'HumanAgent Interface', contains a 'Take Action' section with a text input field showing '{"App.css": "body: { display: none }"}' and a blue 'Step' button. Below this are two buttons: 'Reset Environment' and 'Get State'. The 'Current State' section shows 'Status: Running', 'Episode ID: 1dad6a0f-d665-4633-bc67-6d7df3de3cef', and 'Step Count: 1'. The right panel, titled 'State Observer', displays 'Current Observation' as a JSON object. Below it, there is a large block of JavaScript code for a search results page, including a search function and a 'No results found' message. The bottom of the right panel shows 'Reward: 0 DONE'.