AI and Machine Learning for Testers

Jason Arbon, CEO @Appdiff
Relevant Context

Bing: Testing Neural Net Ranker

Google: Personalized Web Search and Chrome Test Automation

APPDIFF: AI for Mobile Test Automation
Agenda

AI For Testing
Testing AI
Future
The Real Data Scientists

Joanne Tseng
Data Scientist @appdiff.com

Francis Iannacci
Lead Data Scientist @appdiff.com

Appdiff mission:
Transform app development with automation & insights
Testing: Apply Inputs, checking outputs.
AI: Apply Inputs, check outputs

Testing Needs AI

Features

Complexity increases exponentially as new features and states interact with existing features.

Tests

Test coverage grows linearly because tests can only be added one at a time.
Definition of ARTIFICIAL INTELLIGENCE

1. a branch of computer science dealing with the simulation of intelligent behavior in computers

2. the capability of a machine to imitate intelligent human behavior
When will AI Start Testing?
HOW A BOT LEARNS TO TAP, TYPE, AND SWIPE

PERCEIVE
“What can we interact with?”

ACT
Neural network determines what action to take

LEARN
Each action contributes to training brain

[Diagram showing the interaction between an app and a neural network with input, hidden, and output layers, illustrating the process of perceiving, acting, and learning.]
ML: Subjective or complex labels via Humans

Would you like to receive push notifications to be notified when it's time to check in, when your room is ready and be the first to receive exclusive mobile app offers?

Not Now  OK

Choose a category:

- Intro Screen/Welcome Screen
- Login Screen/Sign Up Screen
- Privacy Policy Screen/Terms and Conditions Screen
- Home Screen
- Side Menu Screen
- Popup Screen
- Others

Describe for Others Category:
AI for Testing: Input

Reduce input space

150 actions per page

35 steps -> $150^{35} = 10^{78}$ paths

#atoms in universe $10^{78}...$
AI REPLACES HARD-CRAFTED TEST CASES
SORRY FOR THE DELAY
WE HAD SOME TROUBLE CONNECTING, BUT SHOULD HAVE YOU MOVING SHORTLY.

TRY AGAIN

OR SIGN OUT
EXAMPLE: HOW TO DETECT A BUG

App structure: Small
UI elements: Alert box
Complexity: Low
Text: “Sorry for the delay”
“Try again”

Classification: BUG
AI Driven Test Flows
AI Driven Test Flows

Flow 0

STEP# 0

STEP# 1

Memory: 5MB
CPU: 1%
AI Driven Test Flows
Coverage
<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bot Steps</td>
<td>117,533</td>
</tr>
<tr>
<td>Performance Tests</td>
<td>2,700</td>
</tr>
<tr>
<td>All Actions</td>
<td>391</td>
</tr>
<tr>
<td>Unique Elements</td>
<td>2,455</td>
</tr>
<tr>
<td>Pages</td>
<td>488</td>
</tr>
<tr>
<td>Compute Hours</td>
<td>30</td>
</tr>
</tbody>
</table>
ML: App Changes Don’t Break Bot Testing!
AI bots can test almost any app. 10,000+ apps tested.
Testing the Conference App
<table>
<thead>
<tr>
<th>Speed Category</th>
<th>Benchmark 1</th>
<th>Benchmark 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fastest Screen</td>
<td>Current Build 283ms</td>
<td>Apps in Shopping 192ms</td>
</tr>
<tr>
<td>Average Screen</td>
<td>Current Build 986ms</td>
<td>Apps in Shopping 388ms</td>
</tr>
<tr>
<td>Slowest Screen</td>
<td>Current Build 3044ms</td>
<td>Apps in Shopping 2675ms</td>
</tr>
</tbody>
</table>
Portfolio

Apps in Portfolio

Portfolio Average Stats

Portfolio Ranking Apps By Metrics

Non Crashes

Avg Timing

Max Timing
Current AI Coverage: ~2/3rds Existing Tests

- Long Sequences of Dependant Actions and Verifications
- Basic Tasks (Login, Search, Create Account, Add items to Card, etc.)
- Specific Sequences of Events with Specific Input (search for ‘beanie babies’, etc.)
Automation Coverage: Bots soon 100%
Automation Coverage: Bots... soon after, 10X

Canonical Tests

Learning Tests Cross-app
Testing AI: Test Graph Generation and Traversal
Testing Flow with AI
Abstract Intent Language: AIT

Focus on Intent, not “how”

Human Readable

App-Independant
HARNESSING COLLECTIVE INTELLIGENCE FOR TESTING
Appdiff Represents the Evolution of Software Quality

1. AD HOC TESTING
   Reactively test

2. MANUAL TESTING
   Proactively test

3. TEST AUTOMATION
   Automate repetition

4. AI-DRIVEN APPROACH
   Accelerate coverage