High Performance Machine Learning for The Data Driven Enterprise

MARTIN HACK
CEO & Co-Founder
martin@skytree.net
What’s next?

Past:
- Basic Analytics

Present:
- Advanced Analytics
  - Reports
  - Queries
  - Alerts
  - Human

Future:
- Auto-ML
  - AI Automation
  - Self-tuning Algorithms
  - Massive Computational Search
    - Machine|Human
  - Human|Machine
  - Optimization
  - Predictive
  - Statistical
  - Exploratory

SKYTREE
Scale: Multiple Machines for Higher Performance

Strong scaling*

Data:
- 64 nodes (1024 cores)

*Skytree also supports weak scaling

<table>
<thead>
<tr>
<th>Nodes ( N )</th>
<th>Cores</th>
<th>Speedup (ideal)</th>
<th>Mean Speedup ( \mu ) (Skytree Server)</th>
<th>Standard Deviation ( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>1</td>
<td>1.00013</td>
<td>0.0122826</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>2</td>
<td>2.00665</td>
<td>0.00644421</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>4</td>
<td>3.95269</td>
<td>0.00432337</td>
</tr>
<tr>
<td>8</td>
<td>128</td>
<td>8</td>
<td>7.9136</td>
<td>0.0296349</td>
</tr>
<tr>
<td>16</td>
<td>256</td>
<td>16</td>
<td>15.6988</td>
<td>0.128573</td>
</tr>
<tr>
<td>32</td>
<td>512</td>
<td>32</td>
<td>31.2751</td>
<td>0.110139</td>
</tr>
<tr>
<td>64</td>
<td>1024</td>
<td>64</td>
<td>61.1329</td>
<td>0.867193</td>
</tr>
</tbody>
</table>
Speed – Scale – Accuracy = Competitive Advantage

Micro Targeting – Find New Customers

- **Global 100 Financial Institution**
- Major Pain points: Speed & Accuracy of Current approach

**CURRENT:**
- 1,200 Cores @ 100 Node Hadoop Cluster
- Runtime: 100 Minutes
- Accuracy (Gini): 57%

**SKYTREE SERVER:**
- 12 Cores @ 1 Node
- 1250x speedup
- Runtime: 8 Minutes
- Accuracy (Gini): 60%
Skytree has invented ways to reduce the complexity of ML methods from $O(N^2)$ and $O(N^3)$ to $O(N)$ or $O(N \log N)$.
SAME DATA. BETTER RESULTS.

Thank you.