microprediction™

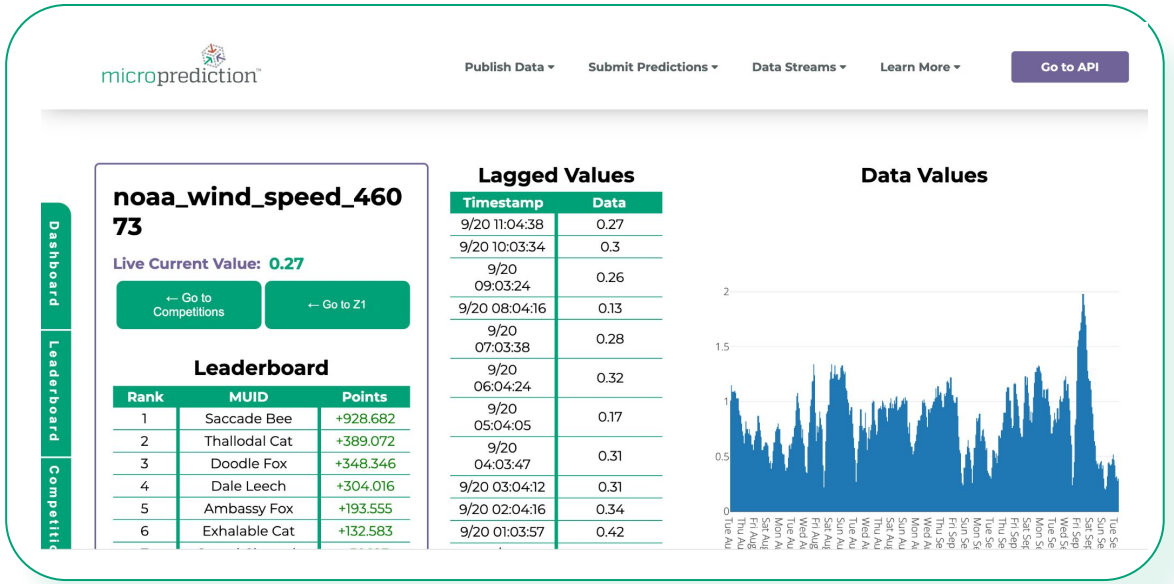
# If You Can Measure It, Consider it Predicted

# HOW IT WORKS

microprediction™

1. You instrument ***anything***

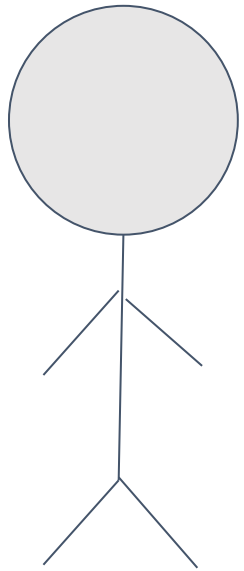
**2.** You [publish](#) numbers so that our system can initiate …



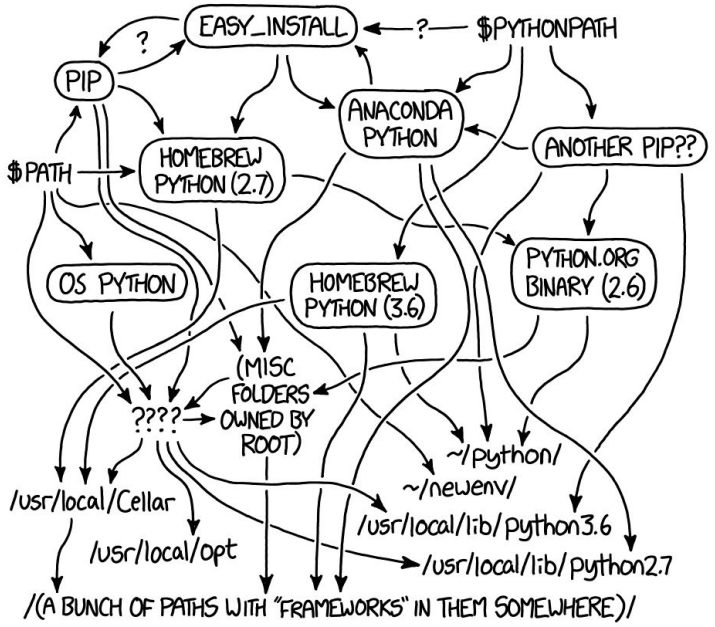

**3.** A **live,** open contest between **algorithms** authored by **anyone,** anywhere

yielding **ongoing** [short-term forecasts](#)

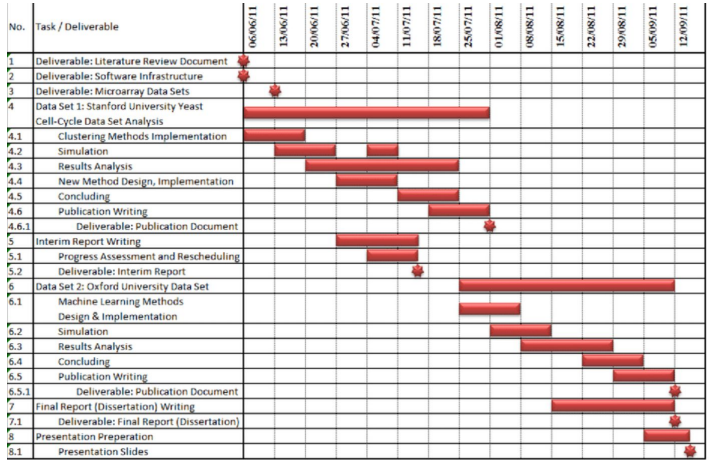
# THE ALTERNATIVE



+




+




Expensive people

Technology bloat

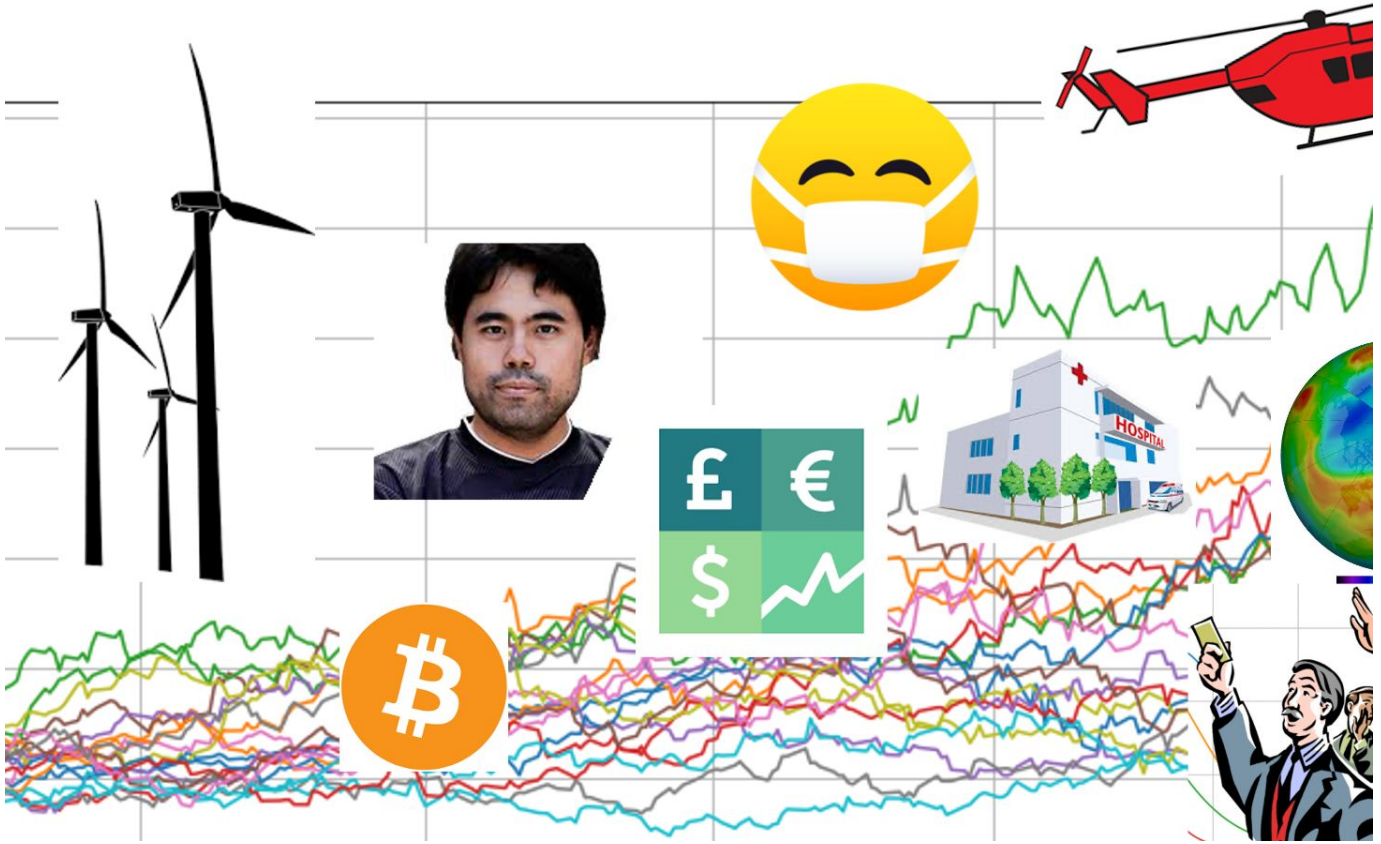
Strung out projects

In M5, 93% of data scientists underperformed exponential smoothing (post)

# SIMPLE PRICING MODEL

microprediction™

$10 PER STREAM / YEAR

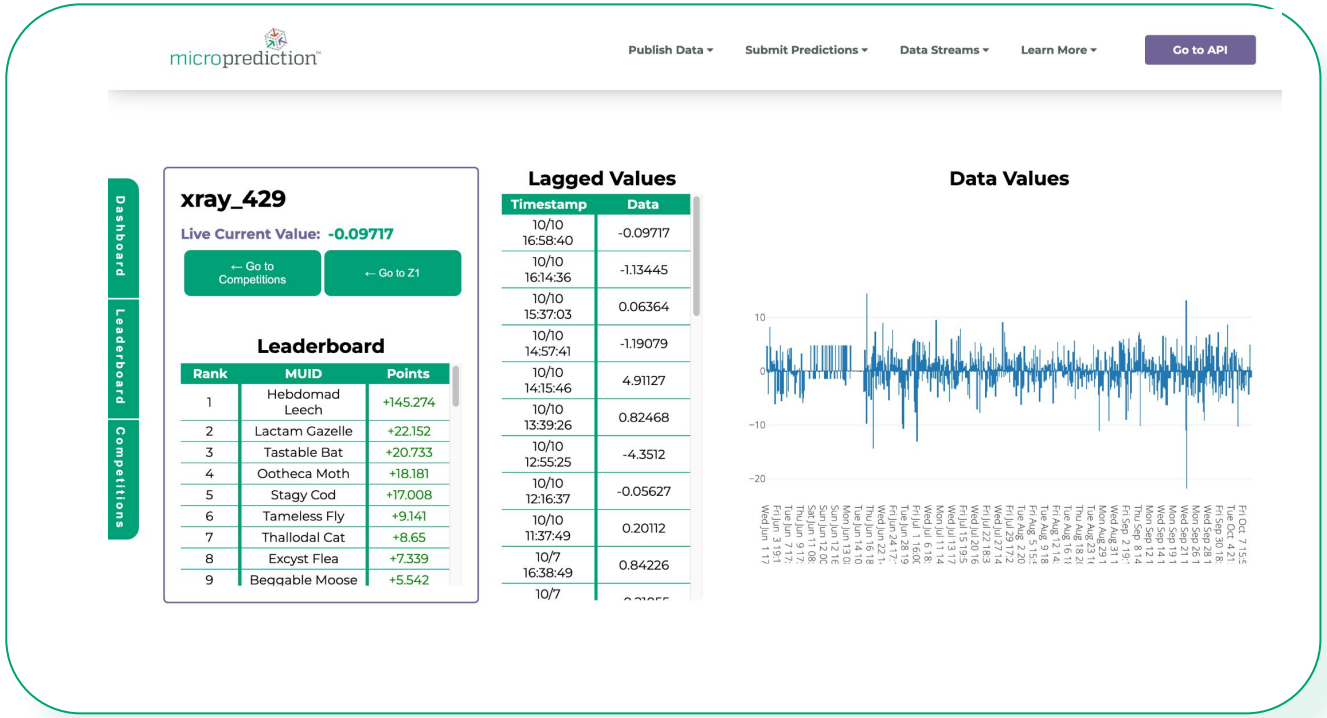



$5 PRIZE-MONEY
$2 TECHNOLOGY COSTS
**$3 PROFIT / STREAM / YR**

# ANCHOR CUSTOMER: INTECH

Using **6000 streams** now

Likely to ramp to 60000 streams




See [live stream example] and [live one hour ahead predictions]

# HOUSE ALGORITHMS ENSURE QUALITY

microprediction™

Private and open source house algorithms informed by ongoing live benchmarking of all popular Python packages

More accurate

Less accurate

| Name | Rating | Games | Active | Seconds | Dependencies |
|---|---|---|---|---|---|
| tsa_precision_combined_ensemble | 2100.0 | 27 | yes | 583.1 | statsmodels , timemachines |
| tsa_p2_d0_q1 | 2082.0 | 399 | yes | 172.9 | statsmodels , timemachines |
| orbit_lgt_24 | 2049.0 | 14 | yes | 51.7 | orbit-ml , timemachines |
| tsa_p1_d0_q1 | 2018.0 | 380 | yes | 107.0 | statsmodels , timemachines |
| bats_damped_arma | 2018.0 | 19 | yes | 1089.1 | tbats , timemachines |
| sk_ae_add_damped | 1990.0 | 1084 | yes | 11.4 | sktime , timemachines |
| tsa_precision_d0_ensemble | 1984.0 | 41 | yes | 282.1 | statsmodels , timemachines |
| sk_autoarima | 1974.0 | 72 | yes | 151.4 | sktime , timemachines |
| | | | | | |
| elo_fastest_univariate_balanced_ensemble | 1719.0 | 1470 | yes | 0.7 | timemachines |
| nprophet_p2_hypocratic | 1712.0 | 449 | yes | 36.7 | neuralprophet , timemachines |
| elo_fastest_residual_aggressive_ensemble | 1709.0 | 1806 | yes | 2.7 | timemachines |
| bats_arma | 1704.0 | 12 | yes | 252.7 | tbats , timemachines |
| tsa_aggressive_theta_ensemble | 1685.0 | 1180 | yes | 3.5 | statsmodels , timemachines |
| pycrt_median_8 | 1674.0 | 4 | yes | 2879.6 | pycaret , timemachines |
| | | | | | |
| fbprophet_univariate | 1598.0 | 238 | yes | 154.8 | prophet , timemachines |

## Classic

- Ensembles of classical models or extensions like TBATS, re-fit and re-calibrated every data point at great computational cost, are the most reliable across diverse time series.

## Machine Learning

- Includes noble attempts to use neural networks such as neuralprophet (results).

## Popular

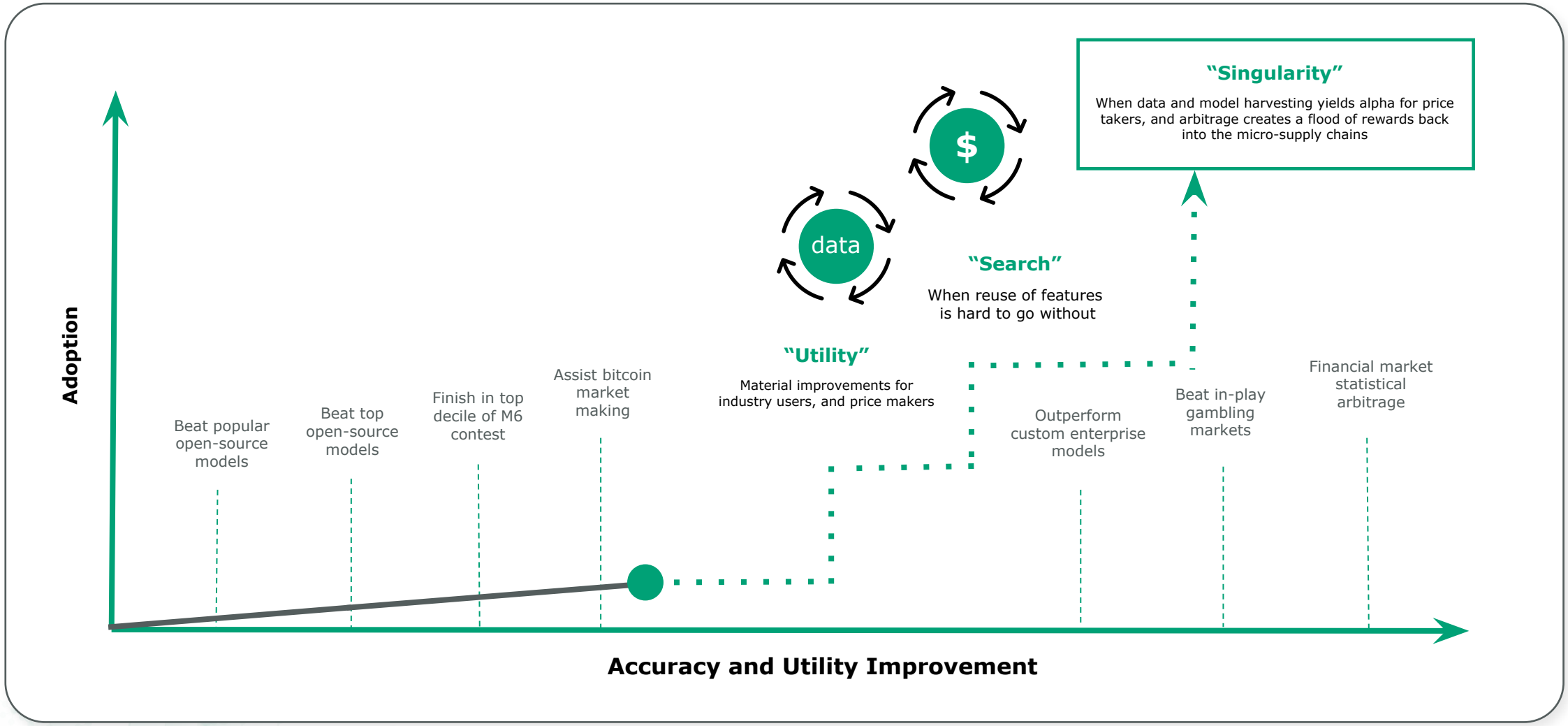
- Facebook Prophet has been downloaded twice as often as any other package. But see our article.

# COMPETITION: PAAS

| | Microprediction | AWS Forecast AutoBox | Explorium PredictHQ | Causalens, DataRobot | BlueMix, Predix |
|---|---|---|---|---|---|
| Starts immediately | ✓ | X | X | X | X |
| Software free | ✓ | X | X | X | X |
| Accurate | ✓ | ? | ? | ? | ? |
| Brings in data | ✓ | X | ✓ | X | X |
| Continuously improves | ✓ | X | ? | ? | X |
| **Eventually unbeatable** | ✓ | X | X | X | X |

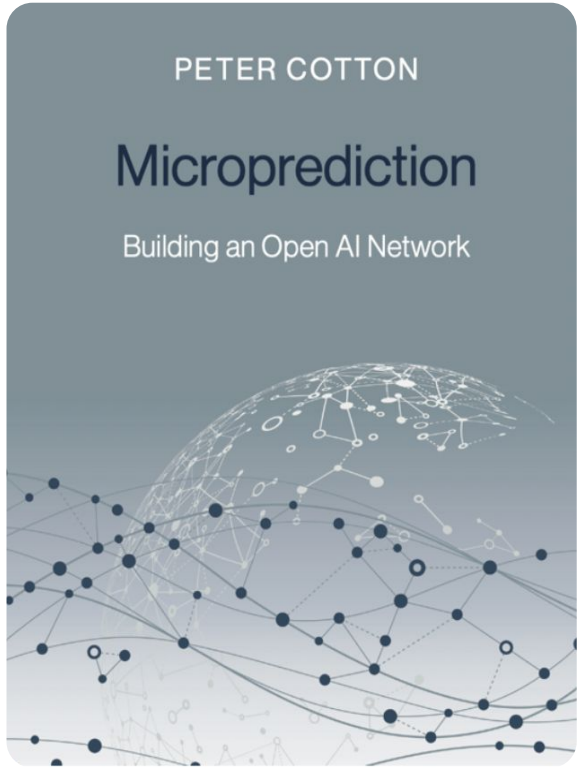
# EVENTUALLY UNBEATABLE

ACCURACY OF THE PREDICTION NETWORK GETS BETTER ALL THE TIME



# TAM EXAMPLES ~ $20b

| Direct | Count | Each ('000s) | Total (billions) |
|---|---|---|---|
| Hedge funds | 4000 | 500 | 2 |
| Licenced brokers | 600000 | 2.5 | 1.5 |
| Transport companies | 1000000 | 5 | 5 |
| Manufacturing companies | 700000 | 5 | 3.5 |
| Agribusiness | 2700000 | 1 | 2.7 |
| Brokers | 600000 | 1 | 0.6 |
| | | | 15.3 |
| **OEM** | | | |
| CRM providers | 2000 | 2000 | 4 |
| IoT platform providers | 600 | 2,000 | 1.2 |
| Alternative data vendors | 400 | 1000 | 0.4 |
| | | | 5.6 |




**Sector survey in book**

**Artificial Intelligence as a Service**

AI-AAS **$92b** by 2030, CAGR of 39%
Precedence Research, 2021 (link)

**Automated Machine Learning**

AutoML **$15b** by 2030, CAGR of 45%
ResearchAndMarkets.com, Feb '22 (link)

# NEEDED IN EVERY SUB-SUB-SECTOR

microprediction™

From Chapter 2 of [Microprediction: Building An Open AI Network]

| Category | Example Sub-cat. | Example Sub-sub cat. |
|---|---|---|
| Recognition | Image | Facial |
| Search | Travel | Personalization |
| Recommendation | Ad-tech | Click-throughs |
| Government | Open cities | Flight status |
| Sales and CRM | Repeat shopping | Visitation |
| Internet of things | Homes | Usage |
| Environment | Air | Pollutants |
| Transport | Driving | Distracted driver |
| Manufacturing | Industrial control | Predictive maintenance |
| Agriculture | Juice | Orange juice |
| Finance | Investment banking | Commercial loans |
| Energy | Power | Wind |
| Medicine | Inventory | Hospital stays |

Table 2.1: Fanout in taxonomy of categories of applications for crowd-sourced microprediction. For each major category we list only one sub-category. For each sub-category I list only one sub-sub-subcategory.

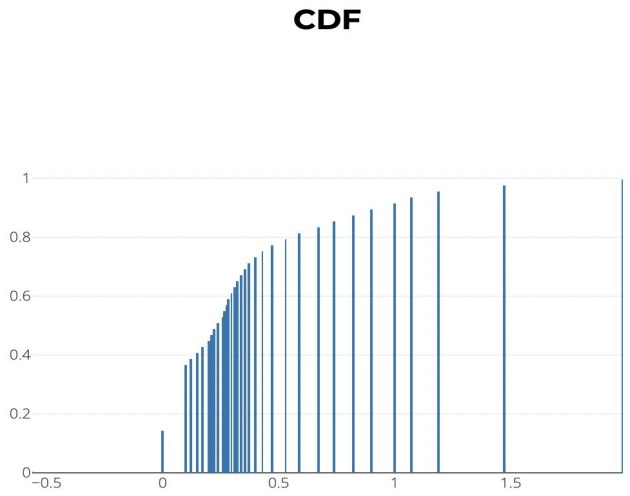
# EXAMPLE USAGE PATTERN

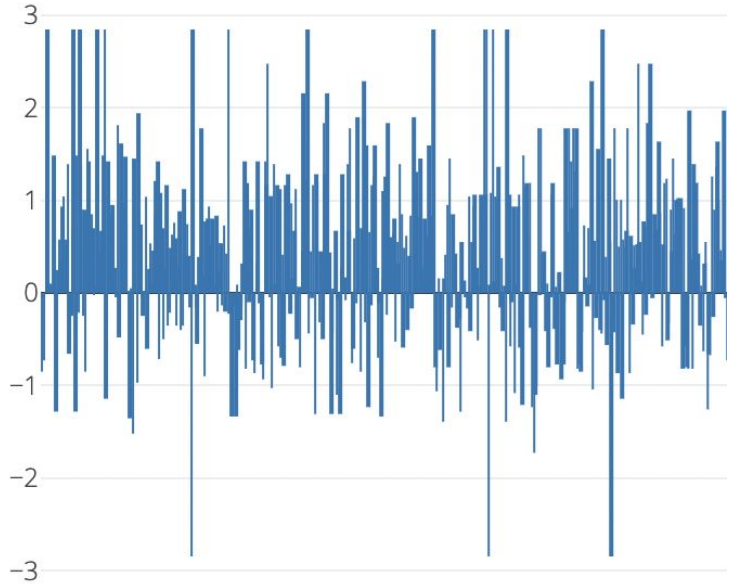
## Business need:

Ongoing performance analysis of in-house models

**Lagged Values**

| Timestamp | Data |
|---|---|
| 9/20 11:04:38 | 0.27 |
| 9/20 10:03:34 | 0.3 |
| 9/20 09:03:24 | 0.26 |
| 9/20 08:04:16 | 0.13 |
| 9/20 07:03:38 | 0.28 |
| 9/20 06:04:24 | 0.32 |
| 9/20 05:04:05 | 0.17 |
| 9/20 04:03:47 | 0.31 |
| 9/20 03:04:12 | 0.31 |
| 9/20 02:04:16 | 0.34 |
| 9/20 01:03:57 | 0.42 |




## Immediate solution:

Send model errors

From Chapter 9 of [Microprediction: Building An Open AI Network]

# NON-FINANCIAL EXAMPLE

microprediction™

Live wind speed

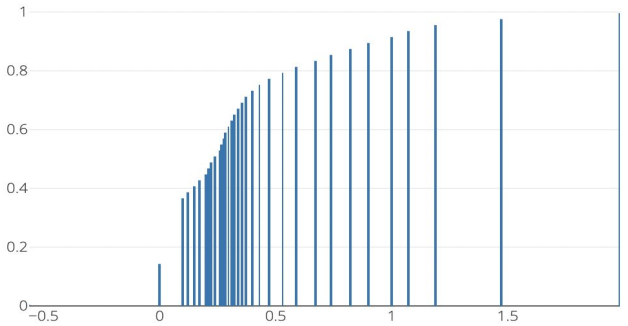
## Actionable intelligence:

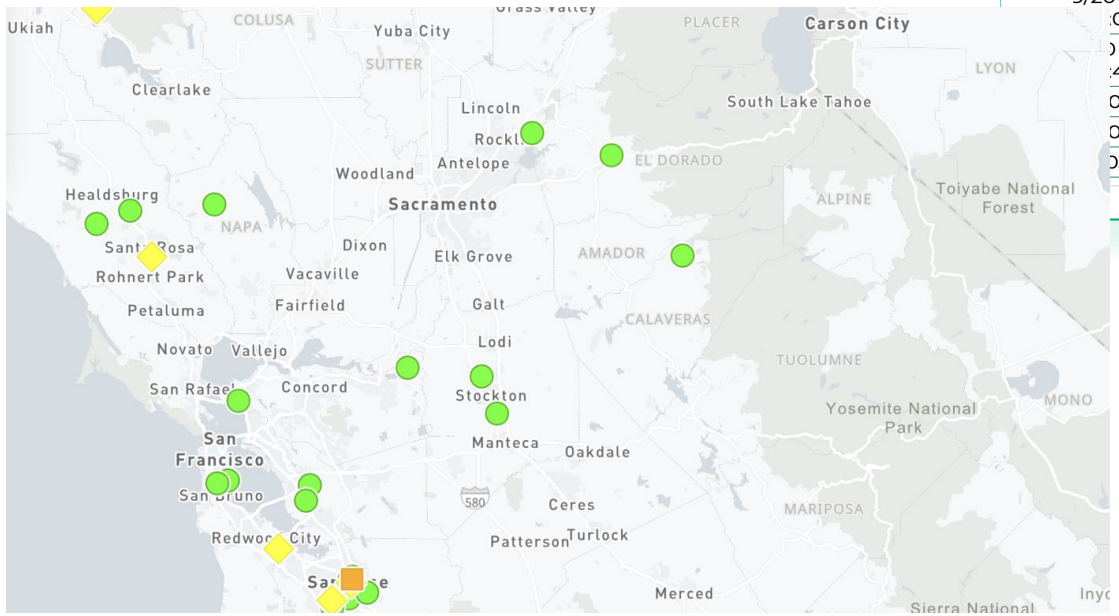
Chance of wind-speed exceeding 40 mph one hour hence

**Lagged Values**

| Timestamp | Data |
|---|---|
| 9/20 11:04:38 | 0.27 |
| 9/20 10:03:34 | 0.3 |
| 9/20 09:03:24 | 0.26 |
| 9/20 08:04:16 | 0.13 |
| 9/20 07:03:38 | 0.28 |
| 9/20 06:04:24 | 0.32 |
| 9/20 :05 | 0.17 |
| :47 | 0.31 |
| 04:12 | 0.31 |
| 04:16 | 0.34 |
| 03:57 | 0.42 |

**CDF**




## Decision:

Implement Public Safety Power Shutoff, potentially costing billions.

See our article "Shutting Down California - The Billion Dollar Prediction Problem"

# ADDITIONAL PRODUCTS and CHANNELS

microprediction™

Enhance any data stream. Improve any prediction.

## Enterprise Data, API Aggregators

- CloudQuant (agreed), Snowflake marketplace, Open FACTSET, Thomson Reuters marketplace, RapidAPI, Quandl, Dawex, Bloomberg, Refinitiv, Informatica, IBM, Oracle, Azure ML service, IDC

## Prediction-AAS OEM Partnerships

- Amazon TS
- Google AI
- Azure ML studio
- IBM

## Early adopters

- Energy
- Finance (Intech, Market-makers, Fraud detection)
- Industrial applications and IoT
- Cyber
- Ad Tech

**See also:**

https://alternativedata.org/data-providers/

# PROGRESS

## Stable platform

… at www.microprediction.org has operated smoothly for over a year, handled a billion predictions in aggregate

## Benchmarking

… of dozens of popular open-source packages and Elo ratings

## Insight

Intech investments discovered novel covariance estimation techniques

## Performance

The two users of precise are 2nd, 19th out of 186 teams in a worldwide contest

## Mindshare

- Rapid ascent to top tier DS commentator
- 20,000+ LI 2,000 new in Sep '22
- ~1,000 in slack

# ABOUT PETER COTTON

## Entrepreneur

- Founded Julius Finance in 2007, raising 1M seed capital.
- Renamed "Benchmark Solutions" with 25M Series A in 2009.
- Invented and wrote a scalable system for autonomous, real-time, curved-based pricing of bonds and credit default swaps en masse.
- Sold FIX data feeds to buy and sell side credit market participants.
- Received $86M verbal to buy company from IDC.
- Technology sold to Bloomberg and integrated into the terminal as BBG:BMRK.

## Leader, inventor and builder

- Lead Morgan Stanley's quantitative credit analytics effort 2002-2006, having identified shortcomings in existing risk analytics for CDOs, and convinced the firm to address them.
- Responsible for numerous new approaches to trading adopted by J.P. Morgan: the use of control theory, data science platforms, microstructure inference and privacy preserving computation.
- Author of Python packages for time-series, optimization, online estimation, structure learning in keras. Approximately 500,000 downloads (GitHub/microprediction).
- Diverse theoretical contributions in partial differential equations, time-series analysis, sports analytics, quantitative finance, epidemiology, statistics and so forth (GitHub/microprediction/home). Author of 8 US Patents.
- Creator and maintainer of a live exchange at microprediction.org that has processed approximately one billion predictions.

PhD, Mathematics, Stanford University - Australian Junior Chess Champion - Medalist in International Physics and Australian Mathematics Olympiads.

# INFLUENCE

microprediction™

## Recommended as the #1 content provider … above **Yann LeCun**


Blog | 12 min

Is Facebook's "Prophet" the Time-Series Messiah, or Just a Very Naughty Boy?

Published on February 3, 2021

**Richad Nieves-Becker** • 1st

Data science leader helping new DS leaders lead with joy & impact | Sr. …

1w •

There are only 4 data science people on LI that I drop everything for to read their content.

Why? They blow my mind and help me solve my problems.

Follow them:

1. **Peter Cotton .Pdf**

The creator of Microprediction. The Prometheus of forecasting. Reminds me how much I don't know.

His hot takes are too clever for the plebs. I love it.

His takedown of Facebook Prophet is legendary. Reminds me of the raptor in Jurassic Park. "Clever girl..."

Oh yeah - he wrote a book you should buy on open prediction networks.

2. **Yann LeCun**

One of the creators of the convolutional net and Chief AI Scientist at Meta.

At least half his content blows my mind. One of the best at framing problems on the frontiers of machine learning.

Reminds me that deep learning isn't just blind empiricism and expensive compute.

The best researchers are philosophers first and ask the right questions.

3. **Cassie Kozyrkov**

# SUMMARY: VALUE PROPOSITION

FINISH THE **DATA SCIENCE PROJECT** WHILE SITTING THROUGH THE FIRST MEETING

# SUMMARY: FUNDING GOALS

microprediction™

1

**Raising $5m**

- Generic prediction API and clients.
- Providing autonomous prediction of rapidly changing instrumented processes, and data feed enhancements.
- Anticipated burn $2-3m.  7 technical, 2 non-tech ramping; prize-money, hardware, services, rent, overhead.

2

**Scale:** 3rd party estimates USD:

- AutoML addressable market $15b by 2030, 45% CAGR (source) with 60% adopting (source)
- AI-AAS $92 Bn by 2030, at a CAGR of 39%. (source)
- Global Alt-data $143b 2030, CAGR 54% (source)
- IoT $2.5T., CAGR 26% (source)

3

**Moats:**

- Serve as central registrar
- Data feedback
- Algorithm feedback
- Preferred enterprise microprediction oracle

microprediction™

# THANK YOU!