
Google™
11
I/O

Smart App Design
with the Google Prediction API

Travis Green, Robert Kaplow, Max Lin
May 11th, 2011

Hashtags: **#io2011 #GoogleAPIs**

Feedback: **goo.gl/N7kRY**

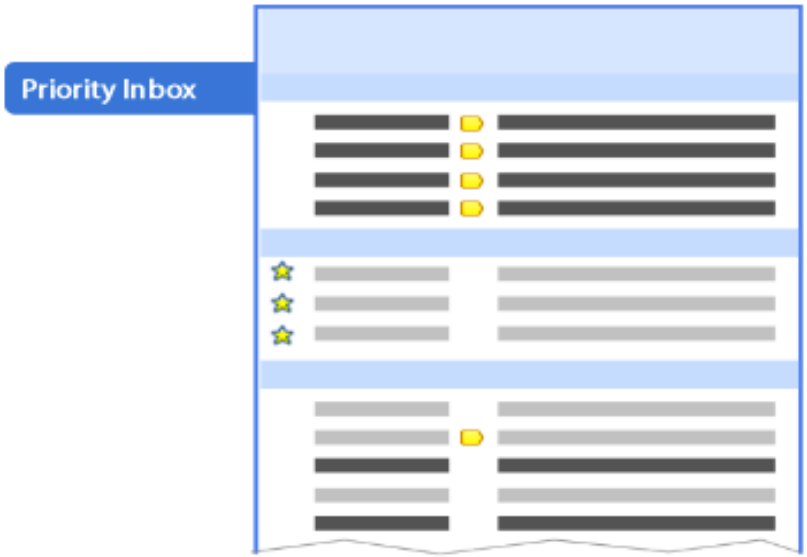
What is a smart app?

- Automates the repetitive

OR

- Recommends the useful

- Extracts the essential



- Pandora
- Cabsense
- ....

What is machine learning?

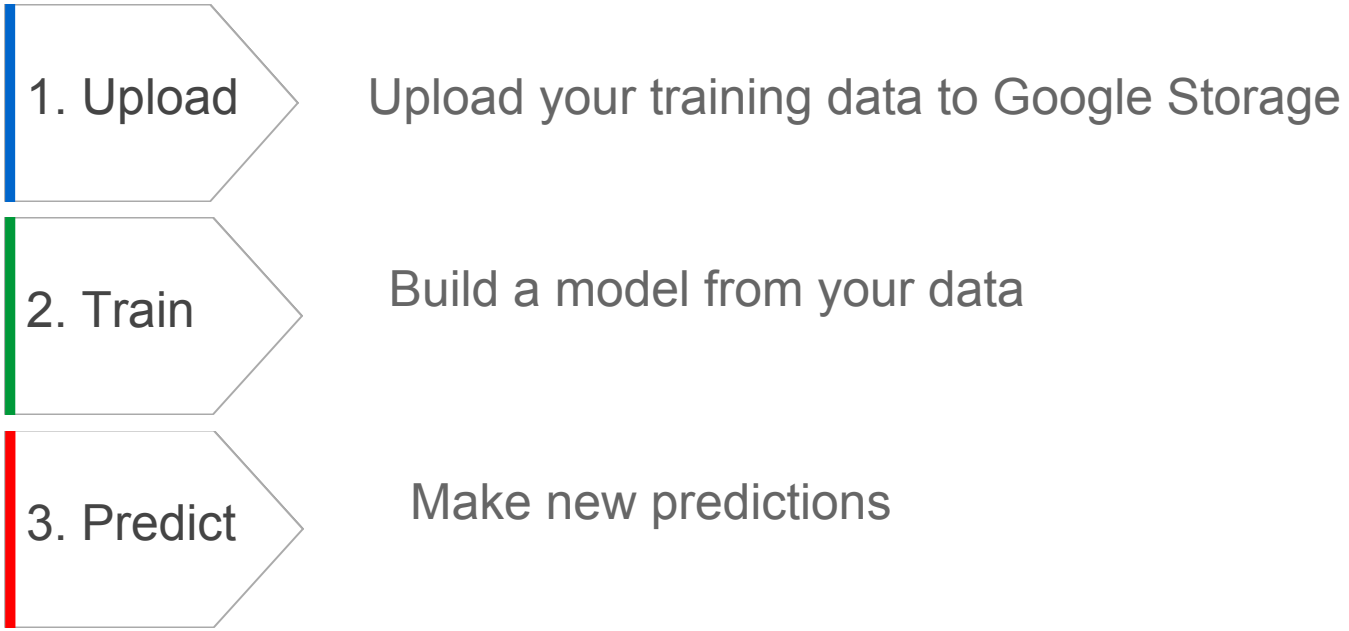
- A set of algorithms that learn patterns from data and make intelligent decisions

Inputs → Predictive Model → **Output**

How do I build one? (Prediction API 101)

- Makes it easy to develop smart apps that automatically respond to your streams of data

- Steps to a smart app:



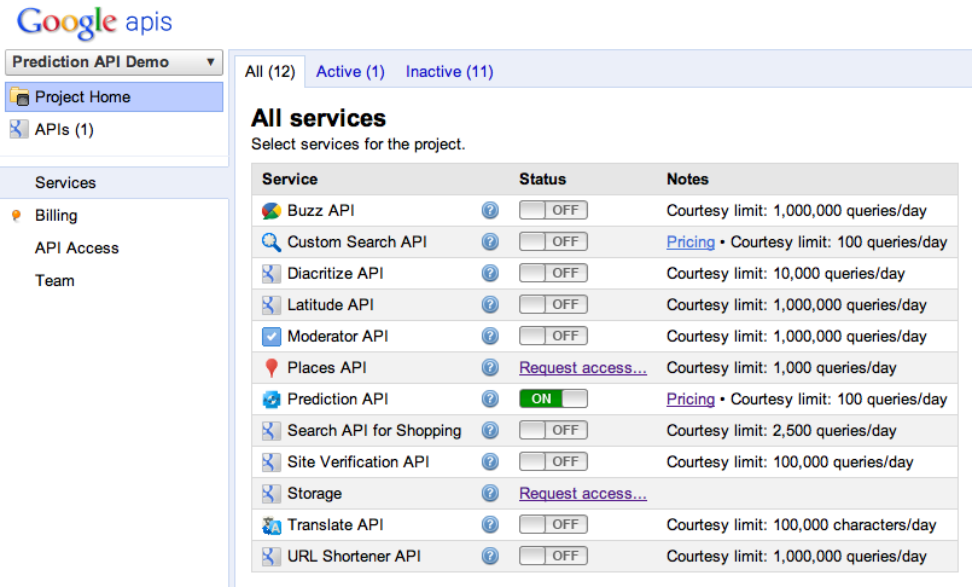
Some Announcements

- Prediction API is generally available
- 99.9% service level agreement for paid users
- Stay tuned!

Building a smart app

Step 0: Signup

Through APIs Console



Google APIs Console

http://code.google.com/apis/console

Step 1: Upload

Upload your training data to Google Storage

- Training data: outputs and input features
- Data format: comma separated value (CSV)

```
$ head -n 2 my_data
"#appengine","The Python library for hosted SQL..."
"#android","I love my Nexus S."
```

- Upload to Google Storage

```
$ gsutil cp my_data gs://io11/my_data
```

Step 2: Train

Create a new model by training on data

- To train a model:

POST prediction/v1.2/training
{"id": "io11/my_data"}

```
{"kind": "prediction#training",
 "id": "io11/my_data",
 "selfLink":"https://www.googleapis.com/prediction/v1.2/training/io11%2Fmy_data"
}
```

- Check training status:

GET prediction/v1.2/training/io11%2Fmy_data

```
{
"kind":"prediction#training",

 "id":"io11/my_data",
 "selfLink":"https://www.googleapis.com/prediction/v1.2/training/io11%2Fmy_data",
  "modelInfo":{
    "modelType":"classification",
    "classificationAccuracy":0.99}
```

Google APIs Explorer

http://code.google.com/apis/explorer

Step 3: Predict

Apply the trained model to make predictions on new data

- To predict:

```
POST prediction/v1.2/training/io11%2Fmy_data/predict
{"input":{
  "csvInstance":["What's a smart Android app?"]}
```

```
{
"kind":"prediction#output",

  "id":"io11/my_data",
  "selfLink":"https://www.googleapis.com/v1.2/training/io11%2Fmy_data/predict",
  "outputLabel": "#android",
  "outputMulti":[{"label":"#android", "score":0.xx},
              {"label":"#appengine", "score":0.xx},
              ...]}
```

[OR - for regression]

```
{
"kind":"prediction#output"

  "id":"io11/my_data",
  "selfLink":"https://www.googleapis.com/v1.2/training/
io11%2Fmy_data/predict",
```

Ford Research

How can I add more data?

Streaming Training

Add real-time data to your predictive model

- New API feature

- Advantages:
  - Adapt quickly to new data
  - Automatically improve performance over time
  - Alternate way to train predictive models

Step 4: Adapt

Stream new data to your predictive model

```
PUT prediction/v1.2/training/io11%2Fmy_data
{
 "classLabel":"#youtube",
 "csvInstance":["I just saw this wicked cool video"]
}

 {"kind":"prediction#training"
  "id":"io11/my_data"
  "selfLink:"https://www.googleapis.com/prediction/v1.2/training/io11%2Fmy_data"
}
```

What if I need data?

Hosted Model Subscriptions

- Users can subscribe to others' models
- Hosted model revenue shared with model developer
- Advantages to users
  - Data already gathered and labeled
  - Built by experts
  - Easy to add to your app

Hosted Model Subscriptions

- Access any model just like a normal prediction
- All models are already enabled for all users

```
POST prediction/v1.2/hostedmodels/model_name/predict
{"input":{
  "csvInstance":[col1_value, col2_value, ...]}
}
```

- What demo models are available?
  - sample.sentiment
  - sample.tagger
  - sample.languageid

Import predictive models

- Predict with models you've already made
- Import models via open-standard Predictive Model Markup Language (PMML)
- Previewing import via a subset of PMML features

Applications

Customer Sentiment

Transaction Risk

Recommend Products

Message Routing

Diagnostics

Churn Prediction

Legal Docket Classification

Suspicious Activity

Work Roster Assignment

Inappropriate Content

Email Filtering

Political Bias

Uplift Marketing

Species Identification

Career Counselling

... and many more ...

Conclusions

- Announcements
  - Prediction API generally available with paid SLA
  - Streaming training data
  - Gallery of hosted predictive models
  - Import your models through PMML (preview)

Questions?

Get Started!

- Try the API: **goo.gl/Jh3jS**

http://code.google.com/apis/predict

Hashtags: **#io2011 #GoogleAPIs**

Feedback: **goo.gl/N7kRY**