

Machine Learning Group

- Motivation
- Current Work
- Scope
- Next Steps

Motivation

- Background: Hal Varian (Google Chief Economist) talk on Google Data Tools
 - *Google Trends*
 - *Google Coorelate*
 - Bayesian Structured Time Series
- ML everywhere
- Potential applications



Scope

- Interest Group
 - a space where we share our expertise and interests
- Outreach
 - Making the library ready
- Systems
 - Making library systems ready

Interest Group

- Theme: Academic side of machine learning
- Model: ACM Special Interest Groups
- Activities: Sharing papers, in-depth exploration of algorithms, etc.
- Intent: team readiness, professional development, complementing individual learning
- Deliverables: Publications, talk sessions, etc.

Outreach

- Theme: Getting the library ready
 - Model: Linux user groups (there are groups that hold “*hack fests*” and “*install fests*”). MIT Media Lab meeting has something similar. Fan clubs.
 - Activities: Engaging the library community; partnering with the code club and the Media Lab; building demos on corpus to brainstorm innovative discovery interfaces
 - Deliverables: New discovery visualizations, etc.
-



Software

- Theme: Product development; intelligent software
- Model: Google Data Tools
- Activities: doing code hackathons; working as a team on different projects
- Deliverables: architecture to enable ML; a suite of ML-powered programs, tools for the broader community, etc.
- Goal: ML compatibility for systems (as outlined in Chris Bourg's blog post)

Current Work

- Word2vec: word similarity
- Classifying search queries in Bento
- Topic model over DSpace abstracts, etc.

Next Steps

- Finishing the set of prototypes
- Identifying project proposals
- Website (currently a blog), spreading the word
- Exploring collaboration with other teams
- Formalizing the set up, creating a charter, etc.