**Domain Areas (Layer 1)**

Government Operation (4), Commercial (8), Defense (3), Healthcare and Life Sciences (10), Deep Learning and Social Media (6), The Ecosystem for Research (4), Astronomy and Physics (5); Earth, Environmental and Polar Science (10) and Energy (1).

**Notes on below**

General Support Tools

Common Alg Components

General Machine Learning

Application Patterns

Application Specialization has 2 levels

A) Used in several Applications

Information Retrieval (text of all sorts, Images) // Image Processing (Pattern Recognition) // Bioinformatics // HPC simulation results // Spatial analysis

B) Then we finally get

Domain1 Domain2 Domain3 .......

So for example

**Tweet analysis, Google News** uses Information Retrieval which uses General Machine Learning // Support Indexing, Search, Query //  Support Streaming

**Language Translation, Voice Recognition** use HMM (Common Methods), Deep Learning

**Driverless cars**use Image Processing,  Deep Learning

The last table goes with Table 3 of Ogre paper to discuss 51 use cases

Support Streaming is a bit vague. It would include data management as in Storm but also could reference streaming algorithms in these were not in General machine Learning

e.g. Streaming clustering is different from Kmeans etc. I don't think there are good streaming libraries -- Mahout does not have?

**Analytics**

**General Support Tools**

Support GPU

Optimize Communication

Support out of memory

**Common Algorithm components**

Nonlinear Solvers (Stochastic Gradient Descent SGD, (L-)BFGS approximation to Newton’s Method, Levenberg-Marquardt solver)

Matrix (SVD, MM multiply, Conjugate Gradient, Eigenvector)

**General Machine Learning**

Basic Statistics

Common Methods (Linear Classifiers, HMM, SVM, Logistic regression, Mixtures, Clustering, outlier, topic models)

Graph

Deep Learning

Recommender Engines

Statistical relational learning

**Application Patterns**

Support Indexing, Search, Query

Support Streaming - Time Series, Sensors

Network and System simulations

Crowd Sourcing

**Application specializations** of above with many areas using

Information Retrieval (text of all sorts, Images)

Image Processing (Pattern Recognition)

Bioinformatics

HPC simulation results

Spatial analysis

Could use Table 2 and 3 of Beowulf paper plus

|  |
| --- |
| Internet of Things (8) and Streaming Apps: Properties and Examples |
| There will be 24 (Mobile Industry Group) to 50 (Cisco) billion devices on the Internet by 2020.  |
| The cloud is natural controller of and resource provider for the Internet of Things. |
| Broad categories are PC’s/tablets/Smart phones/watches, Wearable devices (Smart People), “Intelligent River” “Smart Homes and Grid” “Ubiquitous Cities” and Robotics. |
| Majority of NIST use cases are streaming – experimental science gathers data in a stream – sometimes batched as in a field trip. |
| Use Case | **Title** | **Application** | **Features** |
| 10:  | Cargo Shipping | Tracking as in UPS, Fedex | PP GIS LML |
| 11-13 | Military Sensor Networks | Large Scale Geospatial Analysis and Visualization | PP GIS LML |
| 28 | Truthy | Information diffusion research from Twitter Data | PP MR for Search, GML for community determination |
| 39 | Experimental Particle Physics | Analysis of LHC Large Hadron Collider Data: Discovery of Higgs particle | PP Local Processing Global statistics (MRStat) |
| 50 | Environmental Networks | DOE-BER AmeriFlux and FLUXNET | PP GIS LML |
| 51 | Smart Grids | Energy consumption forecasting | PP GIS LML |

