GENO – GENeric Optimization for Classical Machine Learning

Sören Laue   Matthias Mitterreiter   Joachim Giesen

Friedrich-Schiller-University Jena

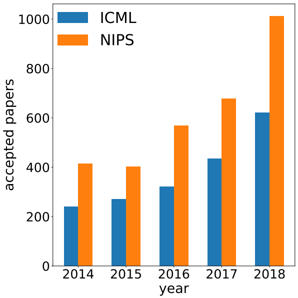
Data Assessment Solutions GmbH, Germany

NeurIPS 2019

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Machine Learning

About 600-2000 papers published at NeurIPS and ICML per year.



~25% design / implement a new optimization algorithm

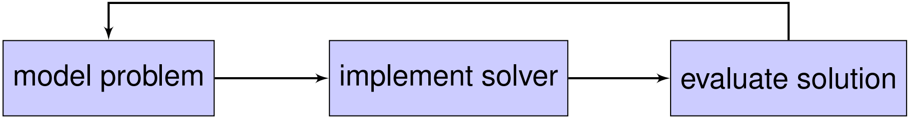
FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Motivation

Look at machine learning through the lens of optimization.

- least squares regression
- logistic regression
- k-means
- network analytics
- (hyperbolic) embeddings
- ...

Work flow



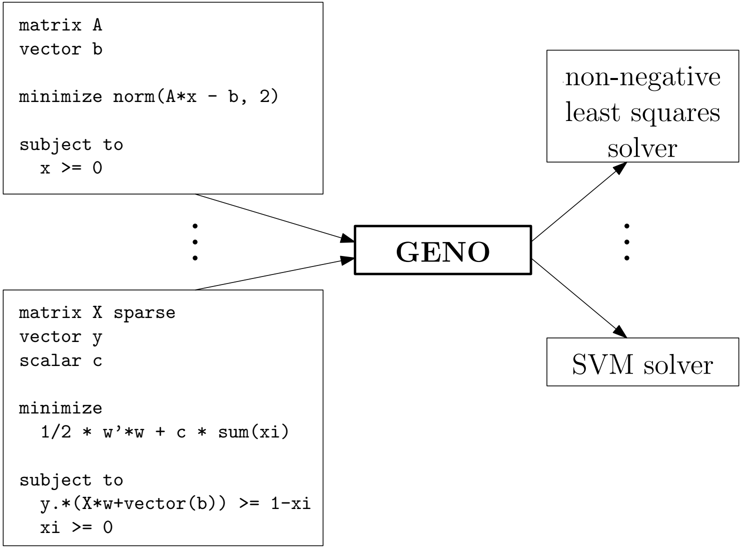
General Optimization in Machine Learning

Ideal world:

- One tool / algorithm for everything
- Easy to use
- As fast as hand-tuned, specialized solvers

Our approach

GENO


matrix A
vector b

minimize norm(A*x - b, 2)

subject to
  x >= 0
GENO
non-negative least squares solver
matrix X sparse
vector y
scalar c

minimize
  1/2 * w'*w + c * sum(xi)

subject to
  y.*(X*w+vector(b)) >= 1-xi
  xi >= 0
SVM solver

Results

Kernelized Dual SVM

data set	LIBSVM	**GENO**	CVX / SeDuMi	CVX / Gurobi
	Sec.	Sec.	Sec.	Sec.
`a1a`	0.28	**0.28**	376.6	57.8
`a7a`	39.5	**29.8**	n/a	n/a

LIBSVM – Chang and Lin 2001

Logistic Regression

data set	LIBLINEAR	**GENO**	CVX / SeDuMi	CVX / Gurobi
	Sec.	Sec.	Sec.	Sec.
`a1a`	0.01	**0.01**	254.3	n/a
`rcv1_test`	12.8	**5.5**	n/a	n/a

LIBLINEAR – Lin et al. (JMLR 2008)

Elastic Net Regression

data set	glmnet Sec.	**GENO** Sec.	CVX/Gurobi Sec.
(1000, 1000)	0.10	**0.11**	21.1
(5000, 10000)	4.75	**4.11**	n/a

glmnet – Friedman, Hastie, Tibshirani (JStatSoft 2010)

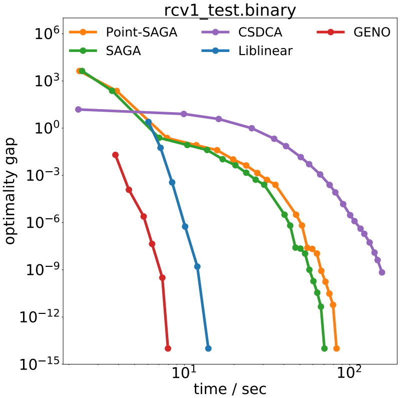
Sparse PCA (non-linear version)

data set	GPower		**GENO**	
	fValue	Sec.	fValue	Sec.
`colon-cancer`	-72.3	0.0109	**-73.6**	**0.0083**
`gisette`	-34.5	1.98	**-34.6**	**1.52**

GPower – Journée et al. (JMLR 2010)

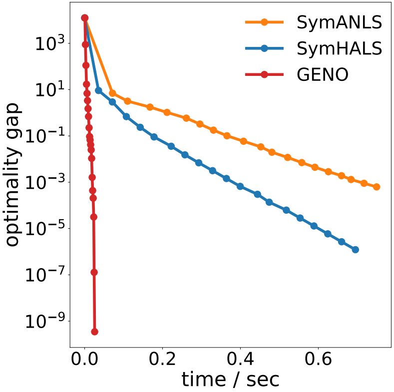
Detailed Results

Logistic Regression



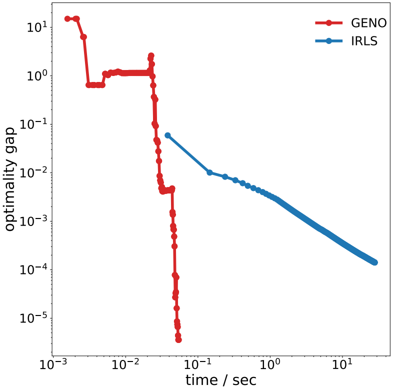
SAGA (NIPS 2014), CSDCA (NIPS 2015), Point-SAGA (NIPS 2016)
LIBLINEAR (JMLR 2008)

Symmetric Non-negative Matrix Factorization



SymANLS, SymHALS (NeurIPS 2018)

Compressed Sensing



IRLS (ICML 2019)

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

GENO

Results:

- One tool / algorithm for ~~everything~~ a lot
- Easy to use
- No tuning needed
- As fast as hand-tuned, specialized, well-established solvers
- Outperforms state-of-the-art, recently published solvers by a large margin
- geno-project.org

Sören Laue, Matthias Mitterreiter, Joachim Giesen.
GENO – GENeric Optimization for Classical Machine Learning. NeurIPS 2019.