

ProDiGe: PRioritization Of Disease Genes with multitask machine learning from positive and unlabeled examples

Fantine Mordelet (CBIO, CREST), Jean-Philippe Vert (CBIO)

(Submitted on 1 Jun 2011)

Elucidating the genetic basis of human diseases is a central goal of genetics and molecular biology. While traditional linkage analysis and modern high-throughput techniques often provide long lists of tens or hundreds of disease gene candidates, the identification of disease genes among the candidates remains time-consuming and expensive. Efficient computational methods are therefore needed to prioritize genes within the list of candidates, by exploiting the wealth of information available about the genes in various databases. Here we propose ProDiGe, a novel algorithm for Prioritization of Disease Genes. ProDiGe implements a novel machine learning strategy based on learning from positive and unlabeled examples, which allows to integrate various sources of information about the genes, to share information about known disease genes across diseases, and to perform genome-wide searches for new disease genes. Experiments on real data show that ProDiGe outperforms state-of-the-art methods for the prioritization of genes in human diseases.

Subjects: **Quantitative Methods (q-bio.QM)**; Machine Learning (stat.ML)

Cite as: [arXiv:1106.0134](https://arxiv.org/abs/1106.0134) [q-bio.QM]

(or [arXiv:1106.0134v1](https://arxiv.org/abs/1106.0134v1) [q-bio.QM] for this version)

Submission history

From: Jean-Philippe Vert [[view email](#)]

[v1] Wed, 1 Jun 2011 09:56:49 GMT (306kb)

Which authors of this paper are endorsers?

Download:

- [PDF](#)
- [PostScript](#)
- [Other formats](#)

Current browse context:

q-bio.QM

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1106](#)

Change to browse by:

[q-bio](#)

[stat](#)

[stat.ML](#)

References & Citations

- [NASA ADS](#)

Bookmark([what is this?](#))



ScienceWISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE

WISE