

# ViperCharts: Visual Performance Evaluation Platform

Borut Sluban and Nada Lavrač

Department of Knowledge Technologies, Jožef Stefan Institute,  
Jamova 39, 1000 Ljubljana, Slovenia  
{borut.sluban,nada.lavrac}@ijs.si

**Abstract.** The paper presents the ViperCharts web-based platform for visual performance evaluation of classification, prediction, and information retrieval algorithms. The platform enables to create interactive charts for easy and intuitive evaluation of performance results. It includes standard visualizations and extends them by offering alternative evaluation methods like  $F$ -isolines, and by establishing relations between corresponding presentations like Precision-Recall and ROC curves. Additionally, the interactive performance charts can be saved, exported to several formats, and shared via unique web addresses. A web API to the service is also available.

**Keywords:** classifiers, performance evaluation, web application

## 1 Introduction and Related Work

Empirical evaluation of classification algorithms is mainly focused on their ability to correctly predict the desired class of data instances. Hence, several performance measures are derived from the confusion matrix, which provides the distribution of the predicted classes over the actual classes of a dataset. The choice of an evaluation measure depends on the task to be solved. Typically, classification is evaluated by accuracy which is the average of correct predictions over all the classes, however in the case of unbalanced class distributions, medical diagnostic or information retrieval tasks other evaluation measures which focus on a certain class, like precision, recall, false positive rate or specificity, may be more desirable. This paper presents the ViperCharts web-based platform for visual performance evaluation of classification, prediction, or information retrieval algorithms used in machine learning and data/text mining. Our goal is to provide a web environment which enables intuitive visualization of results and sharing of performance charts that are of interest to machine learning and data mining practitioners.

Tools for graphical representation of numerical data are provided by different software applications for calculation, numerical computing, statistical computing and statistical analysis, like Microsoft Excel<sup>1</sup>, MATLAB<sup>2</sup>, R<sup>3</sup> and SPSS<sup>4</sup>.

<sup>1</sup> <http://office.microsoft.com/en-us/excel/>

<sup>2</sup> <http://www.mathworks.com/products/matlab/>

<sup>3</sup> <http://www.r-project.org/>

<sup>4</sup> <http://www.ibm.com/software/analytics/spss/>

Also several JavaScript, SVG or Flash based charting libraries and charting software, such as Highcharts<sup>5</sup>, D3<sup>6</sup> or Google Chart Tools<sup>7</sup>, offer data visualizations which can be embedded in web pages. However, among the usually supported chart types only specific line and scatter charts are interesting from the point of visualizing algorithm performance, since they depict the relation between different variables and/or performance measures of algorithms.

Common visualizations of algorithm performance in machine learning include the Lift curves [11], ROC curves [5], Precision-Recall Curves [8], and Cost curves [4], for probabilistic classifiers or ranking algorithms, and scatter charts in the ROC space or Precision-Recall space (PR space) for discrete classification algorithms. Existing environments for machine learning and data mining, like Weka [6], RapidMiner [9], KNIME [1] and Orange [3], as well as MATLAB and R, only offer support for the computation and visualization of ROC and Lift curves, whereas support for other performance visualizations may be available as third-party packages for different programming languages.

In this paper we present the ViperCharts platform, which aims to provide charting support for the evaluation of machine learning, data/text mining and information retrieval algorithms. It offers a range of performance visualizations and reveals the relations among different performance measures. Its main distinctive feature is its web-based design which requires no installation on the user's system, enables sharing of performance results, and offers enhanced visual performance evaluation through interactive charts and additional advanced functionality. Furthermore, it provides a unique environment for computing, visualizing and comparing performance results of different algorithms for various evaluation measures. The ViperCharts platform is accessible online at <http://viper.ijs.si>.

## 2 The ViperCharts Platform

The platform is designed to serve as a tool for data mining practitioners with the goal to produce visualizations of performance results achieved by their algorithms. ViperCharts is a web application running in the client's Web browser. The server side is written in Python<sup>8</sup> and uses the Django Web Framework<sup>9</sup>.

Creating a chart requires two simple steps. First, the user selects the desired chart type, choosing among Scatter chart, Curve chart and Column chart. Where Scatter charts visualize the performance of discrete prediction algorithms in the ROC or PR space, Curve charts visualize Lift, ROC, PR and Cost curves, whereas Column charts visualize arbitrary values for a selection of algorithms. Second, the required data for the specific chart is copy-pasted into the provided form. Different data input formats are supported: TAB delimited data for copy-pasting from spreadsheet applications, CSV data and JSON formatted data. Finally, a *Draw chart* button triggers the chart visualization.

<sup>5</sup> <http://www.highcharts.com>

<sup>7</sup> <https://developers.google.com/chart/>

<sup>9</sup> <http://www.djangoproject.com>

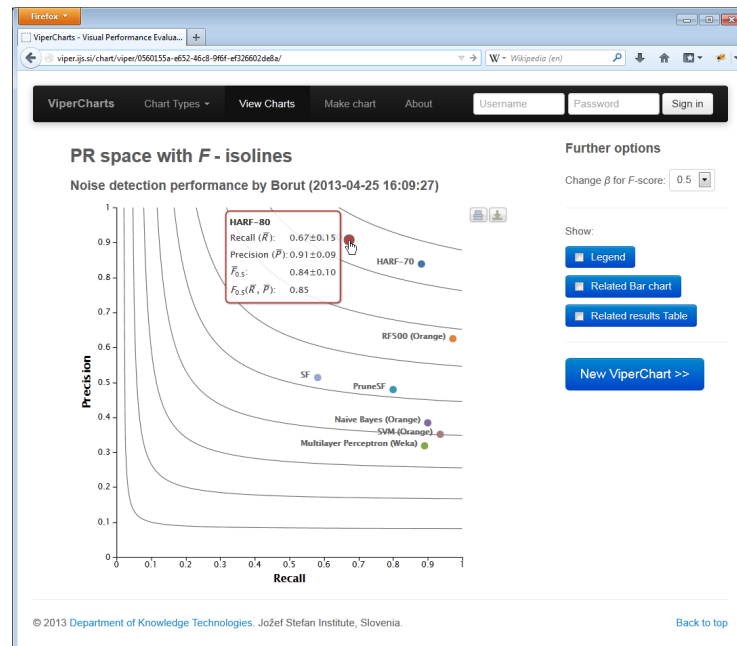
<sup>6</sup> <http://d3js.org>

<sup>8</sup> <http://www.python.org/>

The desired performance measures are calculated from the provided data and stored in the platforms' database. In this way the visualization components have access to the data of a specific chart whenever the chart needs to be displayed.

Charts are drawn with the help of a JavaScript charting library, called Highcharts<sup>5</sup>. For each chart type we created a specific template including individual functionality available for a certain type of performance visualization. For example, in PR space charts we included the novel *F-isoline* evaluation approach [10] which enables to simultaneously visually evaluate algorithm performance in terms of recall, precision and the *F*-measure. As additional novelty, for ROC curve charts a corresponding PR curve chart can be created (and vice versa), since PR curves give a more informative picture of the algorithm's performance when dealing with highly skewed datasets, which provides additional insight for algorithms design, as discussed in [2]. A screenshot of the ViperCharts platform showing a chart of the PR space with *F*-isolines can be found in Figure 1.

The ViperCharts platform enables users to make their charts public and share them with research colleagues, export and publish them in research papers, or include them in their web sites. The sharing and embedding of charts is made easy by a unique URL address assigned to each chart. Exporting the charts into SVG, JPG, PNG or PDF formats is taken care of by the exporting module of the charting library that we use for chart visualization. A web API provides direct access to our performance visualization services, which were also integrated in the cloud based data mining platform, called ClowdFlows [7].



**Fig. 1.** A screenshot of the ViperCharts platform showing the performance of discrete prediction algorithms in the PR space enhanced by *F*-isolines.

### 3 Conclusion and Further Work

We have developed a novel web-based platform for visual performance evaluation of classification, prediction and information retrieval algorithms. The web-based design of the platform allows users to create, save and export their results' visualizations, without any installation needed. Users can easily publish and share their charts, since public charts can be simply accessed through unique web addresses. The performance visualizations can also be accessed through a web API or used as integrated components in the ClowdFlows data mining platform. We are working on covering even a wider range of performance visualizations used by machine learning practitioners, as well as including support for statistical significance tests. We will continue the development on the support for different input data formats and on the relations and conversions between different performance visualizations. We also plan to extend our API to enable integration into several existing data mining environments.

**Acknowledgments.** This work was partially funded by the Slovenian Research Agency and by the European Commission in the context of the FP7 project FIRST under the grant agreement n. 257928.

### References

1. Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Thiel, K., Wiswedel, B.: KNIME - the Konstanz information miner: version 2.0 and beyond. *SIGKDD Explor. Newsl.* 11(1), 26–31 (2009)
2. Davis, J., Goadrich, M.: The relationship between Precision-Recall and ROC curves. In: *Proceedings of the 23rd ICML conference*. pp. 233–240 (2006)
3. Demšar, J., Zupan, B., Leban, G., Curk, T.: Orange: From experimental machine learning to interactive data mining. In: *Proc. of PKDD 2004*, pp. 537–539 (2004)
4. Drummond, C., Holte, R.C.: Cost curves: An improved method for visualizing classifier performance. *Machine Learning* 65(1), 95–130 (2006)
5. Flach, P.A.: ROC Analysis. In: *Encyclopedia of Machine Learning*, pp. 869–875 (2010)
6. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software. *SIGKDD Explorations Newsletter* 11(1), 10–18 (2009)
7. Kranjc, J., Podpečan, V., Lavrač, N.: ClowdFlows: A Cloud Based Scientific Workflow Platform. In: *Proceedings of ECML PKDD 2012*, pp. 816–819 (2012)
8. Manning, C.D., Raghavan, P., Schtze, H.: *Introduction to Information Retrieval*. Cambridge University Press (2008)
9. Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., Euler, T.: YALE: Rapid prototyping for complex data mining tasks. In: *ACM SIGKDD*. pp. 935–940 (2006)
10. Sluban, B., Gamberger, D., Lavrač, N.: Ensemble-based noise detection: noise ranking and visual performance evaluation. *Data Mining and Knowledge Discovery* pp. 1–39 (2013)
11. Witten, I.H., Frank, E.: *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Morgan Kaufmann Publishers Inc. (2005)