The undersampling ensemble approach to highly imbalanced data classification

Paweł Ksieniewicz PAWEL.KSIENIEWICZ@PWR.EDU.PL
Department of Systems and Computer Networks
Faculty of Electronics
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland.

Editor: Editor's name

Abstract

This is the abstract for this article.
Keywords: classification, classifier ensemble, undersampling, imbalanced data

1. Introduction

This is a sample article that uses the jmlr class with the wcp class option. Please follow the guidelines in this sample document as it can help to reduce complications when combining the articles into a book. Please avoid using obsolete commands, such as \rm, and obsolete packages, such as epsfig.[1]

Please also ensure that your document will compile with PDFLaTeX. If you have an error message that's puzzling you, first check for it at the UK TUG FAQ http://www.tex.ac.uk/cgi-bin/texfaq2html?label=man-latex. If that doesn't help, create a minimal working example (see http://theoval.cmp.uea.ac.uk/~nlct/latex/minexample/) and post to somewhere like TeX on StackExchange (http://tex.stackexchange.com/) or the LaTeX Community Forum (http://www.latex-community.org/forum/).

NOTE:
This is an numbered theorem-like environment that was defined in this document's preamble.

1.1. Sub-sections

Sub-sections are produced using \subsection.

1.1.1. Sub-sub-sections

Sub-sub-sections are produced using \subsubsection.

Sub-sub-sub-sections Sub-sub-sub-sections are produced using \paragraph. These are unnumbered with a running head.

1. See http://www.ctan.org/pkg/l2tabu

© 2010 P. Ksieniewicz.

Sub-sub-sub-sub-sections Sub-sub-sub-sub-sections are produced using `\subparagraph`. These are unnumbered with a running head.

2. Cross-Referencing

Always use `\label` and `\ref` (or one of the commands described below) when cross-referencing. For example, the next section is Section 3. The jmlr class provides some convenient cross-referencing commands: `\sectionref`, `\equationref`, `\tableref`, `\figureref`, `\algorithmref`, `\theoremref`, `\lemmaref`, `\remarkref`, `\corollaryref`, `\definitionref`, `\conjectureref`, `\axiomref`, `\exampleref` and `\appendixref`. The argument of these commands may either be a single label or a comma-separated list of labels. Examples:

Referencing sections: Section 3 or Sections 1 and 3 or Sections 1, 3, 5.1 and 5.2.

Referencing equations: Equation (1) or Equations (1) and (3) or Equations (1), (2), (3) and (4).

Referencing tables: Table 1 or Tables 1 and 2 or Tables 1, 2 and 3.

Referencing figures: Figure 1 or Figures 1 and 2 or Figures 1, 2 and 3 or Figures 3(a) and 3(b).

Referencing algorithms: Algorithm 1 or Algorithms 1 and 2 or Algorithms 1, 2 and 3.

Referencing theorem-like environments: Theorem 1, Lemma 2, Remark 3, Corollary 4, Definition 5, Conjecture 6, Axiom 7 and Example 1.

Referencing appendices: Appendix A or Appendices A and B.

3. Equations

The jmlr class loads the amsmath package, so you can use any of the commands and environments defined there. (See the amsmath documentation for further details.[2])

Unnumbered single-lined equations should be displayed using `\[` and `\]`. For example:

$$E = mc^2$$

Numbered single-line equations should be displayed using the `equation` environment. For example:

$$\cos^2\theta + \sin^2\theta \equiv 1 \tag{1}$$

This can be referenced using `\label` and `\equationref`. For example, Equation (1).

Multi-lined numbered equations should be displayed using the `align` environment.[3] For example:

$$f(x) = x^2 + x \tag{2}$$

$$f'(x) = 2x + 1 \tag{3}$$

2. Either `texdoc amsmath` or http://www.ctan.org/pkg/amsmath
3. For reasons why you shouldn't use the obsolete `eqnarray` environment, see Lars Madsen, *Avoid eqnarray!* TUGboat 33(1):21–25, 2012.

Unnumbered multi-lined equations should be displayed using the `align*` environment. For example:

$$\begin{aligned} f(x) &= (x+1)(x-1) \\ &= x^2 - 1 \end{aligned}$$

If you want to mix numbered with unnumbered lines use the align environment and suppress unwanted line numbers with `\nonumber`. For example:

$$\begin{aligned} y &= x^2 + 3x - 2x + 1 \nonumber \\ &= x^2 + x + 1 \tag{4} \end{aligned}$$

An equation that is too long to fit on a single line can be displayed using the `split` environment. Text can be embedded in an equation using `\text` or `\intertext` (as used in Theorem 1). See the amsmath documentation for further details.

3.1. Operator Names

Predefined operator names are listed in Table 1. For additional operators, either use `\operatorname`, for example $\operatorname{var}(X)$ or declare it with `\DeclareMathOperator`, for example

```
\DeclareMathOperator{\var}{var}
```

and then use this new command. If you want limits that go above and below the operator (like `\sum`) use the starred versions (`\operatorname*` or `\DeclareMathOperator*`).

Table 1: Predefined Operator Names (taken from amsmath documentation)

`\arccos`	$\arccos$	`\deg`	$\deg$	`\lg`	$\lg$	`\projlim`	$\projlim$
`\arcsin`	$\arcsin$	`\det`	$\det$	`\lim`	$\lim$	`\sec`	$\sec$
`\arctan`	$\arctan$	`\dim`	$\dim$	`\liminf`	$\liminf$	`\sin`	$\sin$
`\arg`	$\arg$	`\exp`	$\exp$	`\limsup`	$\limsup$	`\sinh`	$\sinh$
`\cos`	$\cos$	`\gcd`	$\gcd$	`\ln`	$\ln$	`\sup`	$\sup$
`\cosh`	$\cosh$	`\hom`	$\hom$	`\log`	$\log$	`\tan`	$\tan$
`\cot`	$\cot$	`\inf`	$\inf$	`\max`	$\max$	`\tanh`	$\tanh$
`\coth`	$\coth$	`\injlim`	$\injlim$	`\min`	$\min$		
`\csc`	$\csc$	`\ker`	$\ker$	`\Pr`	$\Pr$		
		`\varlimsup`	$\varlimsup$	`\varinjlim`	$\varinjlim$		
		`\varliminf`	$\varliminf$	`\varprojlim`	$\varprojlim$		

4. Vectors and Sets

Vectors should be typeset using `\vec`. For example $\boldsymbol{x}$. The jmlr class also provides `\set` to typeset a set. For example $\mathcal{S}$.

5. Floats

Floats, such as figures, tables and algorithms, are moving objects and are supposed to float to the nearest convenient location. Please don't force them to go in a particular place. In general it's best to use the `htbp` specifier and don't put the figure or table in the middle of a paragraph (that is make sure there's a paragraph break above and below the float). Floats are supposed to have a little extra space above and below them to make them stand out from the rest of the text. This extra spacing is put in automatically and shouldn't need modifying.

To ensure consistency, please *don't* try changing the format of the caption by doing something like:

```
\caption{\textit{A Sample Caption.}}
```

or

```
\caption{\em A Sample Caption.}
```

You can, of course, change the font for individual words or phrases, for example:

```
\caption{A Sample Caption With Some \emph{Emphasized Words}.}
```

5.1. Tables

Tables should go in the `table` environment. Within this environment use `\floatconts` (defined by jmlr) to set the caption correctly and center the table contents.

Table 2: An Example Table

Dataset	**Result**
Data1	0.12345
Data2	0.67890
Data3	0.54321
Data4	0.09876

If you want horizontal rules you can use the booktabs package which provides the commands `\toprule`, `\midrule` and `\bottomrule`. For example, see Table 3.

Table 3: A Table With Horizontal Lines

Dataset	**Result**
Data1	0.12345
Data2	0.67890
Data3	0.54321
Data4	0.09876

If you want vertical lines as well, you can't use the booktabs commands as there'll be some unwanted gaps. Instead you can use LaTeX's `\hline`, but the rows may appear a bit cramped. You can add extra space above or below a row using `\abovestrut` and `\belowstrut`. For example, see Table 4.

Table 4: A Table With Horizontal and Vertical Lines

Dataset	**Result**
Data1	0.12345
Data2	0.67890
Data3	0.54321
Data4	0.09876

If you want to align numbers on their decimal point, you can use the siunitx package. For example, see Table 5. For further details see the siunitx documentation[4].

Table 5: A Table With Numbers Aligned on the Decimal Point

Dataset	**Result**
Data1	0.123 45
Data2	10.6789
Data3	50.543
Data4	200.098 76

If the table is too wide, you can adjust the inter-column spacing by changing the value of `\tabcolsep`. For example:

```
\setlength{\tabcolsep}{3pt}
```

If the table is very wide but not very long, you can use the `sidewaystable` environment defined in the rotating package (so use `\usepackage{rotating}`). If the table is too long to fit on a page, you should use the `longtable` environment defined in the longtable package (so use `\usepackage{longtable}`).

5.2. Figures

Figures should go in the `figure` environment. Within this environment, use `\floatconts` to correctly position the caption and center the image. Use `\includegraphics` for external graphics files but omit the file extension. Do not use `\epsfig` or `\psfig`. If you want to scale the image, it's better to use a fraction of the line width rather than an explicit length. For example, see Figure 1.

If your image is made up of LaTeX code (for example, commands provided by the pgf package) you can include it using `\includeteximage` (defined by the jmlr class). This can be scaled and rotated in the same way as `\includegraphics`. For example, see Figure 2.

4. Either `texdoc siunitx` or http://www.ctan.org/pkg/siunitx

Figure 1: Example Image

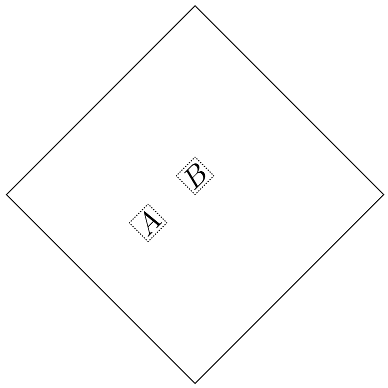


Figure 2: Image Created Using LaTeX Code

If the figure is too wide to fit on the page, you can use the `sidewaysfigure` environment defined in the `rotating` package.

Don't use `\graphicspath`. If the images are contained in a subdirectory, specify this when you include the image, for example `\includegraphics{figures/mypic}`.

5.2.1. Sub-Figures

Sub-figures can be created using `\subfigure`, which is defined by the `jmlr` class. The optional argument allows you to provide a subcaption. The label should be placed in the mandatory argument of `\subfigure`. You can reference the entire figure, for example Figure 3, or you can reference part of the figure using `\figureref`, for example Figure 3(*a*). Alternatively you can reference the subfigure using `\subfigref`, for example (*a*) and (*b*) in Figure 3.

By default, the sub-figures are aligned on the baseline. This can be changed using the second optional argument of `\subfigure`. This may be `t` (top), `c` (centered) or `b` (bottom). For example, the subfigures (*a*) and (*b*) in Figure 4 both have `[c]` as the second optional argument.

5.3. Sub-Tables

There is an analogous command `\subtable` for sub-tables. It has the same syntax as `\subfigure` described above. You can reference the table using `\tableref`, for example

Dataset	Without oversampled set										With oversampled set										OS	US	Full
	All members					Reduced members					All members					Reduced members							
	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC			
ecoli-0-1-3-7-vs-2-6	.845	.843	.843	.837	.837	.845	.845	.845	.837	.837	.841	.841	.839	.828	.825	.845	.845	.845	.837	.837	.806	.838	.825
ecoli4	.809	.834	.859	.879	.857	.748	.800	.820	.926	.898	.874	.898	.868	.913	.904	.746	.896	.895	.923	.922	.859	.787	.878
glass-0-1-6-vs-2	.580	.580	.610	.580	.610	.577	.577	.577	.583	.580	.552	.552	.552	.560	.591	.577	.577	.577	.585	.585	.569	.589	.580
glass-0-1-6-vs-5	.989	.989	.989	.939	.941	.989	.989	.989	.989	.989	.941	.941	.941	.941	.941	.989	.989	.989	.989	.989	.941	.975	.941
glass2	.619	.619	.619	.616	.616	.641	.641	.641	.641	.641	.619	.616	.619	.616	.616	.641	.641	.644	.641	.641	.617	.620	.591
glass4	.800	.797	.800	.802	.802	.807	.807	.804	.766	.766	.781	.781	.774	.779	.779	.812	.774	.774	.771	.771	.731	.745	.587
glass5	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.938	.945	.938
page-blocks-1-3-vs-4	.817	.817	.817	.846	.845	.828	.828	.828	.828	.828	.807	.807	.803	.846	.845	.831	.831	.831	.830	.828	.791	.806	.763
shuttle-c0-vs-c4	.994	.994	.994	.995	.994	.994	.994	.994	.991	.991	.995	.995	.995	.991	.991	.994	.994	.994	.991	.991	.990	.993	.991
shuttle-c2-vs-c4	.971	.975	.975	.979	.979	.996	.996	.996	.996	.996	.984	.979	.979	.988	.988	.996	.996	.996	.996	.996	.988	.946	.996
vowel0	.909	.909	.909	.902	.903	.909	.909	.909	.909	.909	.910	.910	.910	.904	.903	.910	.909	.909	.909	.909	.906	.905	.917
yeast-0-5-6-7-9-vs-4	.690	.706	.689	.728	.731	.680	.737	.727	.786	.774	.702	.700	.677	.726	.715	.687	.724	.710	.785	.791	.498	.601	.504
yeast-1-2-8-9-vs-7	.563	.570	.567	.576	.570	.562	.557	.562	.627	.642	.563	.552	.553	.574	.566	.556	.564	.556	.625	.639	.540	.598	.544
yeast-1-4-5-8-vs-7	.552	.553	.550	.557	.563	.567	.566	.554	.550	.575	.551	.558	.554	.555	.590	.588	.555	.562	.564	.568	.541	.566	.547
yeast-1-vs-7	.698	.698	.699	.705	.703	.703	.710	.689	.726	.722	.671	.681	.674	.705	.700	.700	.695	.674	.725	.719	.586	.686	.604
yeast-2-vs-4	.804	.801	.803	.799	.800	.820	.822	.827	.870	.862	.773	.794	.781	.799	.800	.805	.833	.827	.869	.861	.529	.739	.561
yeast-2-vs-8	.773	.773	.773	.773	.774	.773	.773	.773	.773	.773	.796	.771	.796	.773	.773	.773	.773	.773	.773	.773	.616	.762	.657
yeast4	.675	.674	.674	.769	.728	.708	.766	.766	.822	.813	.644	.655	.637	.781	.746	.710	.765	.769	.820	.811	.526	.660	.551
yeast5	.934	.934	.934	.934	.934	.929	.934	.938	.957	.955	.917	.921	.919	.934	.934	.927	.930	.935	.957	.955	.780	.910	.831
yeast6	.785	.787	.787	.847	.796	.829	.863	.848	.887	.878	.760	.773	.768	.843	.795	.818	.860	.845	.887	.878	.628	.795	.650
ecoli-0-1-4-6-vs-5	.785	.781	.783	.804	.804	.660	.810	.810	.906	.913	.810	.812	.860	.906	.906	.785	.837	.862	.910	.913	.885	.679	.877
ecoli-0-1-4-7-vs-2-3-5-6	.630	.630	.630	.630	.630	.630	.630	.630	.652	.654	.630	.630	.630	.662	.663	.630	.630	.630	.657	.659	.667	.634	.630
ecoli-0-1-4-7-vs-5-6	.617	.617	.617	.710	.693	.617	.617	.617	.823	.829	.637	.677	.755	.852	.855	.617	.617	.617	.826	.831	.863	.668	.735
ecoli-0-1-vs-2-3-5	.578	.578	.578	.575	.578	.558	.578	.578	.658	.658	.618	.618	.638	.638	.638	.558	.618	.618	.658	.658	.639	.578	.638
ecoli-0-1-vs-5	.639	.689	.689	.736	.732	.616	.666	.664	.741	.741	.714	.739	.789	.834	.830	.641	.691	.689	.816	.816	.797	.658	.782
ecoli-0-2-3-4-vs-5	.683	.733	.708	.778	.753	.658	.733	.733	.853	.853	.756	.758	.806	.828	.803	.731	.758	.758	.895	.895	.638	.657	.754
ecoli-0-2-6-7-vs-3-5	.563	.563	.588	.635	.613	.588	.588	.588	.625	.630	.588	.588	.588	.625	.628	.588	.613	.613	.623	.628	.592	.595	.563
ecoli-0-3-4-6-vs-5	.876	.876	.876	.876	.876	.818	.870	.873	.901	.898	.859	.865	.849	.857	.851	.843	.870	.870	.901	.895	.725	.716	.784
ecoli-0-3-4-7-vs-5-6	.633	.633	.633	.765	.729	.633	.673	.653	.781	.799	.697	.735	.726	.785	.747	.653	.673	.673	.789	.791	.734	.665	.775
ecoli-0-3-4-vs-5	.756	.756	.756	.833	.792	.711	.786	.783	.883	.883	.769	.794	.811	.772	.758	.758	.783	.783	.883	.883	.730	.657	.817
ecoli-0-4-6-vs-5	.876	.876	.876	.901	.903	.853	.876	.878	.903	.878	.901	.901	.901	.901	.901	.878	.878	.876	.903	.901	.890	.725	.854
ecoli-0-6-7-vs-3-5	.598	.598	.598	.573	.555	.548	.548	.548	.640	.597	.557	.557	.557	.575	.575	.548	.508	.508	.618	.597	.544	.571	.508
ecoli-0-6-7-vs-5	.795	.787	.793	.805	.812	.685	.755	.705	.853	.830	.845	.838	.838	.828	.825	.688	.755	.755	.853	.830	.847	.682	.780
glass-0-1-4-6-vs-2	.558	.589	.558	.615	.581	.622	.622	.620	.620	.620	.592	.592	.558	.595	.595	.622	.622	.620	.620	.620	.597	.590	.577
glass-0-1-5-vs-2	.567	.582	.582	.527	.534	.542	.536	.542	.604	.582	.527	.527	.527	.530	.533	.533	.551	.558	.584	.584	.508	.555	.519
glass-0-4-vs-5	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.994	.984	.984	.994
glass-0-6-vs-5	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.995	.945	.989	.945
yeast-0-2-5-6-vs-3-7-8-9	.576	.576	.576	.576	.576	.576	.576	.576	.782	.782	.634	.657	.692	.777	.768	.576	.576	.576	.781	.783	.782	.605	.670
yeast-0-2-5-7-9-vs-3-6-8	.883	.893	.889	.901	.901	.901	.901	.902	.897	.900	.779	.834	.841	.900	.898	.894	.896	.896	.897	.900	.524	.785	.577
yeast-0-3-5-9-vs-7-8	.597	.632	.606	.620	.621	.606	.605	.606	.605	.605	.601	.607	.589	.619	.633	.611	.612	.600	.605	.605	.539	.633	.557

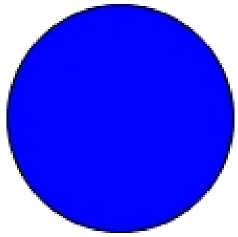
Table 6: Balanced accuracy scores obtained using GNB as a base classifier

Without oversampled set										With oversampled set													
All members					Reduced members					All members					Reduced members								
REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	OS	US	Full	Dataset
.836	.836	.836	.847	.849	.845	.845	.847	.844	.845	.830	.830	.832	.835	.834	.854	.853	.851	.853	.856	.835	.835	.850	*ecoli-0-1-3-7-vs-2-6*
.941	.941	.945	.949	.949	.940	.940	.940	.921	.945	.964	.967	.972	.909	.909	.945	.945	.945	.940	.943	.909	.928	.848	*ecoli4*
.718	.718	.718	.715	.724	.732	.735	.726	.746	.699	.717	.751	.760	.750	.728	.705	.705	.737	.758	.735	.656	.666	.555	*glass-0-1-6-vs-2*
.880	.880	.880	.883	.883	.830	.877	.880	.880	.880	.836	.839	.853	.879	.879	.833	.833	.836	.833	.841	.933	.852	.739	*glass-0-1-6-vs-5*
.696	.696	.724	.721	.719	.701	.724	.718	.724	.724	.737	.757	.746	.756	.756	.726	.732	.731	.723	.746	.715	.678	.485	*glass2*
.861	.861	.861	.878	.875	.861	.861	.861	.863	.863	.868	.868	.888	.905	.913	.861	.861	.861	.888	.898	.925	.865	.781	*glass4*
.823	.823	.826	.830	.830	.813	.813	.813	.813	.816	.838	.843	.879	.801	.873	.816	.813	.826	.821	.833	.830	.811	.695	*glass5*
.868	.868	.874	.877	.877	.867	.871	.866	.902	.897	.901	.907	.926	.909	.924	.874	.875	.881	.917	.920	.917	.872	.808	*page-blocks-1-3-vs-4*
.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	.996	*shuttle-c0-vs-c4*
.859	.859	.859	.891	.859	.879	.883	.850	.883	.850	.996	.000	.000	.000	.000	.887	.887	.906	.927	.996	.000	.845	.600	*shuttle-c2-vs-c4*
.942	.942	.942	.944	.943	.943	.942	.943	.945	.945	.952	.953	.963	.997	.996	.948	.948	.954	.980	.986	.999	.939	.977	*vowel0*
.781	.781	.785	.786	.790	.779	.779	.785	.785	.787	.810	.824	.832	.830	.820	.799	.800	.819	.828	.829	.795	.792	.667	*yeast-0-5-6-7-9-vs-4*
.644	.645	.662	.665	.663	.659	.660	.655	.645	.659	.702	.695	.673	.625	.626	.655	.660	.684	.692	.693	.627	.652	.499	*yeast-1-2-8-9-vs-7*
.576	.574	.597	.593	.592	.571	.571	.602	.571	.584	.622	.628	.614	.612	.609	.594	.607	.623	.636	.643	.615	.590	.499	*yeast-1-4-5-8-vs-7*
.698	.699	.695	.705	.697	.730	.729	.736	.733	.742	.737	.743	.712	.700	.700	.745	.755	.734	.706	.710	.705	.682	.517	*yeast-1-vs-7*
.915	.915	.914	.907	.908	.914	.914	.912	.917	.917	.921	.921	.913	.906	.906	.921	.920	.921	.921	.909	.885	.908	.819	*yeast-2-vs-4*
.715	.715	.740	.719	.719	.721	.721	.728	.727	.734	.782	.782	.778	.803	.803	.738	.738	.752	.755	.758	.803	.734	.774	*yeast-2-vs-8*
.819	.819	.819	.839	.838	.840	.841	.841	.841	.843	.838	.829	.839	.771	.774	.844	.844	.843	.841	.852	.749	.835	.574	*yeast4*
.955	.955	.954	.955	.954	.957	.957	.956	.957	.958	.956	.956	.960	.962	.967	.958	.958	.958	.960	.960	.964	.952	.850	*yeast5*
.883	.883	.884	.886	.885	.886	.886	.885	.887	.886	.899	.900	.892	.840	.865	.887	.887	.888	.893	.895	.840	.879	.739	*yeast6*
.890	.890	.890	.892	.892	.892	.892	.892	.892	.890	.894	.894	.896	.904	.906	.892	.892	.892	.892	.892	.917	.886	.898	*ecoli-0-1-4-6-vs-5*
.888	.888	.888	.890	.885	.893	.895	.891	.896	.896	.869	.869	.886	.869	.886	.886	.886	.889	.904	.902	.856	.882	.847	*ecoli-0-1-4-7-vs-2-3-5-6*
.886	.886	.888	.868	.868	.883	.885	.881	.913	.891	.874	.876	.887	.892	.894	.893	.893	.871	.887	.889	.899	.883	.838	*ecoli-0-1-4-7-vs-5-6*
.892	.892	.892	.914	.894	.892	.892	.872	.892	.872	.872	.865	.867	.885	.885	.874	.874	.872	.883	.863	.887	.895	.830	*ecoli-0-1-vs-2-3-5*
.905	.905	.907	.914	.911	.905	.905	.907	.911	.911	.911	.911	.911	.911	.914	.909	.909	.911	.911	.911	.916	.902	.900	*ecoli-0-1-vs-5*
.900	.900	.900	.895	.897	.900	.900	.903	.900	.903	.900	.900	.900	.900	.900	.900	.900	.900	.900	.900	.909	.904	.894	*ecoli-0-2-3-4-vs-5*
.836	.836	.836	.841	.841	.833	.836	.833	.816	.811	.853	.858	.860	.830	.833	.843	.843	.848	.860	.855	.890	.814	.787	*ecoli-0-2-6-7-vs-3-5*
.887	.887	.874	.890	.890	.887	.887	.879	.890	.890	.893	.893	.895	.901	.901	.893	.893	.890	.895	.895	.911	.881	.875	*ecoli-0-3-4-6-vs-5*
.904	.904	.902	.917	.917	.904	.904	.902	.921	.915	.901	.901	.901	.890	.890	.913	.911	.897	.885	.888	.894	.887	.876	*ecoli-0-3-4-7-vs-5-6*
.903	.903	.903	.903	.900	.900	.900	.903	.903	.900	.906	.906	.906	.903	.903	.906	.906	.906	.903	.903	.911	.889	.875	*ecoli-0-3-4-vs-5*
.890	.890	.890	.887	.884	.884	.884	.884	.890	.890	.890	.890	.890	.898	.895	.890	.890	.890	.892	.895	.914	.889	.900	*ecoli-0-4-6-vs-5*
.860	.860	.860	.872	.875	.857	.857	.855	.872	.867	.875	.875	.855	.865	.863	.865	.870	.870	.863	.858	.893	.844	.835	*ecoli-0-6-7-vs-3-5*
.853	.853	.850	.868	.863	.860	.860	.858	.870	.872	.878	.878	.880	.863	.885	.865	.865	.872	.890	.887	.863	.850	.847	*ecoli-0-6-7-vs-5*
.715	.715	.706	.712	.725	.717	.712	.712	.714	.717	.731	.717	.701	.718	.716	.741	.749	.747	.734	.727	.732	.681	.512	*glass-0-1-4-6-vs-2*
.684	.650	.653	.663	.653	.674	.674	.650	.653	.650	.705	.728	.748	.767	.740	.676	.695	.712	.745	.758	.656	.651	.527	*glass-0-1-5-vs-2*
.932	.932	.939	.932	.932	.939	.939	.944	.939	.944	.951	.951	.951	.951	.951	.951	.951	.951	.951	.951	.988	.917	.850	*glass-0-4-vs-5*
.873	.873	.873	.868	.868	.868	.868	.873	.878	.878	.883	.898	.928	.953	.953	.878	.878	.878	.913	.933	.985	.816	.745	*glass-0-6-vs-5*
.784	.784	.786	.781	.784	.785	.785	.790	.786	.790	.801	.802	.800	.791	.798	.793	.798	.800	.795	.798	.784	.760	.762	*yeast-0-2-5-6-vs-3-7-8-9*
.909	.909	.908	.911	.911	.904	.904	.908	.906	.910	.913	.913	.902	.900	.900	.909	.913	.912	.896	.895	.904	.902	.902	*yeast-0-2-5-7-9-vs-3-6-8*
.757	.757	.755	.746	.753	.748	.748	.757	.749	.754	.759	.742	.739	.731	.730	.770	.767	.750	.723	.734	.718	.702	.639	*yeast-0-3-5-9-vs-7-8*

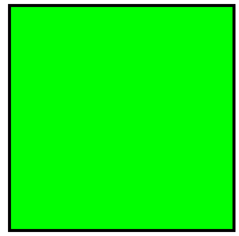
Table 7: Balanced accuracy scores obtained using KNN as a base classifier

Without oversampled set										With oversampled set													
All members					Reduced members					All members					Reduced members								
REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	REG	WEI	CON	NOR	NC	OS	US	Full	Dataset
.816	.816	.816	.818	.816	.823	.823	.823	.838	.838	.783	.784	.784	.715	.725	.823	.823	.823	.842	.842	.624	.708	.841	*ecoli-0-1-3-7-vs-2-6*
.866	.866	.866	.866	.866	.876	.873	.873	.873	.873	.876	.878	.878	.865	.889	.876	.876	.876	.834	.834	.817	.848	.866	*ecoli4*
.740	.737	.731	.657	.657	.732	.727	.727	.680	.680	.716	.716	.722	.603	.584	.741	.738	.738	.700	.700	.581	.630	.546	*glass-0-1-6-vs-2*
.929	.929	.929	.937	.934	.934	.934	.934	.937	.937	.934	.937	.937	.943	.943	.934	.934	.934	.940	.940	.859	.886	.936	*glass-0-1-6-vs-5*
.739	.742	.739	.803	.757	.785	.780	.780	.784	.784	.749	.684	.684	.634	.644	.767	.800	.800	.801	.801	.616	.682	.573	*glass2*
.885	.885	.885	.898	.893	.900	.897	.897	.907	.907	.898	.898	.895	.868	.868	.905	.903	.903	.860	.860	.819	.835	.804	*glass4*
.924	.924	.924	.879	.877	.934	.934	.934	.891	.891	.939	.939	.946	.971	.973	.939	.939	.939	.949	.949	.933	.867	.898	*glass5*
.982	.982	.981	.987	.987	.991	.991	.991	.992	.992	.989	.989	.990	.993	.993	.991	.991	.991	.992	.992	.994	.958	.996	*page-blocks-1-3-vs-4*
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	*shuttle-c0-vs-c4*
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.950	.950	.950	.950	.950	.000	.000	.000	.000	.000	.990	.959	.950	*shuttle-c2-vs-c4*
.955	.955	.955	.956	.956	.957	.956	.956	.956	.956	.959	.959	.958	.951	.951	.957	.957	.957	.961	.961	.921	.940	.936	*vowel0*
.788	.788	.788	.783	.786	.787	.787	.787	.787	.787	.760	.758	.751	.704	.705	.778	.777	.777	.759	.759	.674	.750	.659	*yeast-0-5-6-7-9-vs-4*
.731	.733	.729	.722	.737	.736	.736	.736	.732	.732	.733	.742	.738	.659	.647	.749	.748	.748	.750	.750	.621	.624	.630	*yeast-1-2-8-9-vs-7*
.637	.639	.643	.621	.602	.633	.624	.624	.586	.586	.656	.650	.645	.570	.562	.649	.649	.649	.587	.587	.533	.581	.537	*yeast-1-4-5-8-vs-7*
.767	.768	.764	.784	.781	.783	.794	.794	.801	.801	.737	.741	.709	.622	.598	.782	.782	.782	.789	.789	.603	.661	.683	*yeast-1-vs-7*
.950	.950	.951	.955	.955	.948	.948	.948	.954	.954	.958	.958	.959	.895	.888	.960	.958	.958	.931	.931	.822	.900	.843	*yeast-2-vs-4*
.798	.795	.794	.778	.774	.787	.787	.787	.782	.782	.808	.813	.820	.737	.716	.790	.789	.789	.802	.802	.697	.715	.690	*yeast-2-vs-8*
.824	.824	.817	.837	.826	.846	.846	.846	.844	.844	.815	.818	.830	.651	.642	.850	.850	.850	.845	.845	.626	.792	.643	*yeast4*
.962	.962	.961	.965	.964	.964	.964	.964	.965	.965	.966	.967	.956	.931	.932	.964	.964	.964	.967	.967	.845	.936	.845	*yeast5*
.838	.838	.839	.840	.838	.847	.847	.847	.852	.852	.850	.851	.856	.804	.782	.852	.851	.851	.860	.860	.750	.818	.730	*yeast6*
.900	.900	.898	.873	.873	.910	.910	.910	.887	.887	.885	.885	.885	.871	.871	.892	.887	.887	.860	.860	.794	.823	.781	*ecoli-0-1-4-6-vs-5*
.842	.842	.838	.819	.819	.847	.845	.845	.830	.830	.848	.849	.849	.869	.852	.848	.848	.848	.866	.866	.827	.804	.820	*ecoli-0-1-4-7-vs-2-3-5-6*
.866	.866	.866	.853	.853	.878	.871	.871	.876	.876	.879	.881	.881	.867	.867	.876	.876	.876	.867	.867	.844	.803	.787	*ecoli-0-1-4-7-vs-5-6*
.873	.871	.871	.873	.873	.878	.878	.878	.865	.865	.840	.840	.840	.838	.838	.820	.820	.820	.831	.831	.764	.802	.760	*ecoli-0-1-vs-2-3-5*
.873	.873	.873	.873	.873	.848	.848	.848	.850	.850	.850	.850	.850	.855	.830	.852	.850	.850	.855	.855	.805	.841	.857	*ecoli-0-1-vs-5*
.914	.911	.911	.909	.909	.911	.911	.911	.911	.911	.917	.917	.917	.897	.895	.917	.917	.917	.897	.897	.832	.843	.781	*ecoli-0-2-3-4-vs-5*
.823	.823	.823	.820	.820	.833	.823	.823	.820	.820	.833	.833	.833	.838	.838	.833	.830	.830	.835	.835	.811	.791	.790	*ecoli-0-2-6-7-vs-3-5*
.882	.882	.884	.901	.904	.879	.879	.879	.887	.887	.893	.893	.893	.893	.893	.906	.901	.901	.901	.901	.812	.834	.786	*ecoli-0-3-4-6-vs-5*
.875	.875	.875	.853	.853	.875	.875	.875	.853	.853	.866	.866	.866	.859	.859	.868	.864	.864	.851	.851	.836	.839	.840	*ecoli-0-3-4-7-vs-5-6*
.928	.928	.928	.914	.914	.928	.928	.928	.911	.911	.936	.936	.936	.944	.944	.911	.936	.936	.947	.947	.869	.862	.831	*ecoli-0-3-4-vs-5*
.912	.912	.912	.912	.912	.906	.906	.906	.912	.912	.906	.906	.906	.829	.831	.906	.904	.904	.859	.859	.813	.838	.836	*ecoli-0-4-6-vs-5*
.818	.818	.818	.825	.825	.818	.818	.818	.830	.830	.832	.832	.840	.865	.867	.840	.835	.835	.853	.853	.864	.786	.850	*ecoli-0-6-7-vs-3-5*
.867	.867	.867	.875	.875	.882	.875	.875	.878	.878	.863	.863	.863	.838	.812	.880	.880	.880	.838	.838	.827	.819	.795	*ecoli-0-6-7-vs-5*
.835	.833	.830	.805	.810	.835	.827	.827	.746	.746	.847	.813	.796	.711	.681	.739	.739	.739	.715	.715	.676	.675	.610	*glass-0-1-4-6-vs-2*
.806	.803	.803	.836	.840	.734	.800	.800	.819	.819	.704	.704	.704	.557	.557	.754	.751	.751	.535	.535	.572	.634	.578	*glass-0-1-5-vs-2*
.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.982	.994	.942	.994	*glass-0-4-vs-5*
.959	.959	.959	.974	.974	.959	.959	.959	.980	.980	.964	.964	.969	.985	.985	.975	.975	.975	.990	.990	.955	.879	.995	*glass-0-6-vs-5*
.775	.775	.770	.758	.757	.782	.782	.782	.790	.790	.800	.802	.805	.742	.723	.801	.795	.795	.780	.780	.701	.732	.733	*yeast-0-2-5-6-vs-3-7-8-9*
.903	.903	.903	.888	.887	.898	.898	.898	.895	.895	.896	.896	.896	.876	.878	.903	.900	.900	.897	.897	.867	.868	.854	*yeast-0-2-5-7-9-vs-3-6-8*
.736	.736	.737	.709	.716	.729	.729	.729	.724	.724	.728	.724	.724	.627	.628	.720	.715	.715	.651	.651	.599	.635	.688	*yeast-0-3-5-9-vs-7-8*

Table 8: Balanced accuracy scores obtained using DTC as a base classifier

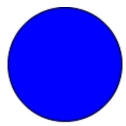


(a) A Circle

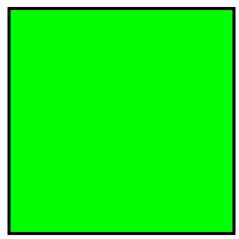


(b) A Square

Figure 3: An Example With Sub-Figures.



(a) A Small Circle



(b) A Square

Figure 4: Another Example With Sub-Figures.

Table 9 or you can reference part of the table, for example Table 9(a). Alternatively you can reference the subtable using `\subtabref`, for example (a) and (b) in Table 9.

Table 9: An Example With Sub-Tables

(a)

A	**B**
1	2

(b)

C	**D**
3	4
5	6

By default, the sub-tables are aligned on the top. This can be changed using the second optional argument of `\subtable`. This may be `t` (top), `c` (centered) or `b` (bottom). For example, the sub-tables (a) and (b) in Table 10 both have `[c]` as the second optional argument.

5.4. Algorithms

Enumerated textual algorithms can be displayed using the `algorithm` environment. Within this environment, use use an `enumerate` or nested `enumerate` environments. For example, see Algorithm 1. Note that algorithms float like figures and tables.

You can use `\caption` and `\label` without using `\floatconts` (as in Algorithm 2).

Table 10: Another Example With Sub-Tables

(a)

A	**B**
1	2

(b)

C	**D**
3	4
5	6

Algorithm 1: The Gauss-Seidel Algorithm

1. For $k = 1$ to maximum number of iterations
   (a) For $i = 1$ to n
      i. $x_i^{(k)} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)}}{a_{ii}}$
      ii. If $\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)} < \epsilon\|$, where ϵ is a specified stopping criteria, stop.

If you'd rather have the same numbering throughout the algorithm but still want the convenient indentation of nested `enumerate` environments, you can use the `enumerate*` environment provided by the jmlr class. For example, see Algorithm 2.

Algorithm 2: Moore's Shortest Path
Given a connected graph G, where the length of each edge is 1:

1. Set the label of vertex s to 0
2. Set $i = 0$
   3. Locate all unlabelled vertices adjacent to a vertex labelled i and label them $i+1$
   4. If vertex t has been labelled,
      the shortest path can be found by backtracking, and the length is given by the label of t.

   otherwise

      increment i and return to step 3

Pseudo code can be displayed using the `algorithm2e` environment. This is defined by the algorithm2e package (which is automatically loaded) so check the algorithm2e documentation for further details.[5] For an example, see Algorithm 3.

6. Description Lists

The jmlr class also provides a description-like environment called `altdescription`. This has an argument that should be the widest label in the list. Compare:

5. Either `texdoc algorithm2e` or http://www.ctan.org/pkg/algorithm2e

Algorithm 3: Computing Net Activation

Input: $x_1, \ldots, x_n, w_1, \ldots, w_n$
Output: y, the net activation
$y \leftarrow 0$;
for $i \leftarrow 1$ **to** n **do**
  $y \leftarrow y + w_i * x_i$;
end

add A method that adds two variables.

differentiate A method that differentiates a function.

with

add	A method that adds two variables.
differentiate	A method that differentiates a function.

7. Theorems, Lemmas etc

The following theorem-like environments are predefined by the `jmlr` class: `theorem`, `example`, `lemma`, `proposition`, `remark`, `corollary`, `definition`, `conjecture` and `axiom`. You can use the `proof` environment to display the proof if need be, as in Theorem 1.

Theorem 1 (Eigenvalue Powers) *If λ is an eigenvalue of $\boldsymbol{B}$ with eigenvector $\boldsymbol{\xi}$, then λ^n is an eigenvalue of $\boldsymbol{B}^n$ with eigenvector $\boldsymbol{\xi}$.*
Proof *Let λ be an eigenvalue of $\boldsymbol{B}$ with eigenvector ξ, then*

$$\boldsymbol{B}\boldsymbol{\xi} = \lambda\boldsymbol{\xi}$$

premultiply by $\boldsymbol{B}$:

$$\begin{aligned} \boldsymbol{B}\boldsymbol{B}\boldsymbol{\xi} &= \boldsymbol{B}\lambda\boldsymbol{\xi} \\ \Rightarrow \boldsymbol{B}^2\boldsymbol{\xi} &= \lambda\boldsymbol{B}\boldsymbol{\xi} \\ &= \lambda\lambda\boldsymbol{\xi} \qquad \textit{since } \boldsymbol{B}\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \\ &= \lambda^2\boldsymbol{\xi} \end{aligned}$$

Therefore true for $n = 2$. Now assume true for $n = k$:

$$\boldsymbol{B}^k\boldsymbol{\xi} = \lambda^k\boldsymbol{\xi}$$

premultiply by $\boldsymbol{B}$:

$$\begin{aligned} \boldsymbol{B}\boldsymbol{B}^k\boldsymbol{\xi} &= \boldsymbol{B}\lambda^k\boldsymbol{\xi} \\ \Rightarrow \boldsymbol{B}^{k+1}\boldsymbol{\xi} &= \lambda^k\boldsymbol{B}\boldsymbol{\xi} \\ &= \lambda^k\lambda\boldsymbol{\xi} \qquad \textit{since } \boldsymbol{B}\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \\ &= \lambda^{k+1}\boldsymbol{\xi} \end{aligned}$$

Therefore true for $n = k + 1$. Therefore, by induction, true for all n. ■

Lemma 2 (A Sample Lemma) *This is a lemma.*

Remark 3 (A Sample Remark) *This is a remark.*

Corollary 4 (A Sample Corollary) *This is a corollary.*

Definition 5 (A Sample Definition) *This is a definition.*

Conjecture 6 (A Sample Conjecture) *This is a conjecture.*

Axiom 7 (A Sample Axiom) *This is an axiom.*

Example 1 (An Example) *This is an example.*

8. Color vs Grayscale

It's helpful if authors supply grayscale versions of their images in the event that the article is to be incorporated into a black and white printed book. With external PDF, PNG or JPG graphic files, you just need to supply a grayscale version of the file. For example, if the file is called `myimage.png`, then the gray version should be `myimage-gray.png` or `myimage-gray.pdf` or `myimage-gray.jpg`. You don't need to modify your code. The jmlr class checks for the existence of the grayscale version if it is print mode (provided you have used `\includegraphics` and haven't specified the file extension).

You can use `\ifprint` to determine which mode you are in. For example, in Figure 1, the purple ellipse represents an input and the yellow ellipse represents an output. Another example: important text!

You can use the class option `gray` to see how the document will appear in gray scale mode. Colored text will automatically be converted to gray scale.

The jmlr class loads the xcolor package, so you can also define your own colors. For example: XYZ.

The xcolor class is loaded with the `x11names` option, so you can use any of the x11 predefined colors (listed in the xcolor documentation[6]).

9. Citations and Bibliography

The jmlr class automatically loads natbib. This sample file has the citations defined in the accompanying BibTeX file `jmlr-sample.bib`. For a parenthetical citation use `\citep`. For example (Guyon and Elisseeff, 2003). For a textual citation use `\citet`. For example Guyon et al. (2007). Both commands may take a comma-separated list, for example Guyon and Elisseeff (2003); Guyon et al. (2007).

These commands have optional arguments and have a starred version. See the natbib documentation for further details.[7]

The bibliography is displayed using `\bibliography`.

---

6. either `texdoc xcolor` or `http://www.ctan.org/pkg/xcolor`
7. Either `texdoc natbib` or `http://www.ctan.org/pkg/natbib`

Acknowledgments

Acknowledgements go here.

References

I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *JMLR*, 3: 1157–1182, March 2003.

I. Guyon, C. Aliferis, and A. Elisseeff. Causal feature selection. Technical report, Clopinet, 2007.

Appendix A. First Appendix

This is the first appendix.

Appendix B. Second Appendix

This is the second appendix.