

000 001 002 003 004 005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

000 001 002 003 004 005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

Paper under double-blind review

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

The abstract paragraph should be indented 1/2 inch (3 picas) on both left and right-hand margins. Use 10 point type, with a vertical spacing of 11 points. The word ABSTRACT must be centered, in small caps, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

ICLR requires electronic submissions, processed by <https://openreview.net/>. See ICLR’s website for more instructions.

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

If your paper is ultimately accepted, the statement `\iclrfinalcopy` should be inserted to adjust the format to the camera ready requirements.

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

The format for the submissions is a variant of the NeurIPS format. Please read carefully the instructions below, and follow them faithfully.

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

Papers to be submitted to ICLR 2025 must be prepared according to the instructions presented here.

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

Authors are required to use the ICLR \LaTeX style files obtainable at the ICLR website. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

The style files for ICLR and other conference information are available online at:

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

<http://www.iclr.cc/>

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

The file `iclr2025_conference.pdf` contains these instructions and illustrates the various formatting requirements your ICLR paper must satisfy. Submissions must be made using \LaTeX and the style files `iclr2025_conference.sty` and `iclr2025_conference.bst` (to be used with \LaTeX 2e). The file `iclr2025_conference.tex` may be used as a “shell” for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own.

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

The formatting instructions contained in these style files are summarized in sections 2, 3, and 4 below.

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long. The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing of 11 points. Times New Roman is the preferred typeface throughout. Paragraphs are separated by 1/2 line space, with no indentation.

005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

Paper title is 17 point, in small caps and left-aligned. All pages should start at 1 inch (6 picas) from the top of the page.

Authors' names are set in boldface, and each name is placed above its corresponding address. The lead author's name is to be listed first, and the co-authors' names are set to follow. Authors sharing the same address can be on the same line.

Please pay special attention to the instructions in section 4 regarding figures, tables, acknowledgments, and references.

There will be a strict upper limit of 10 pages for the main text of the initial submission, with unlimited additional pages for citations.

3 HEADINGS: FIRST LEVEL

First level headings are in small caps, flush left and in point size 12. One line space before the first level heading and 1/2 line space after the first level heading.

3.1 HEADINGS: SECOND LEVEL

Second level headings are in small caps, flush left and in point size 10. One line space before the second level heading and 1/2 line space after the second level heading.

3.1.1 HEADINGS: THIRD LEVEL

Third level headings are in small caps, flush left and in point size 10. One line space before the third level heading and 1/2 line space after the third level heading.

4 CITATIONS, FIGURES, TABLES, REFERENCES

These instructions apply to everyone, regardless of the formatter being used.

4.1 CITATIONS WITHIN THE TEXT

Citations within the text should be based on the `natbib` package and include the authors' last names and year (with the "et al." construct for more than two authors). When the authors or the publication are included in the sentence, the citation should not be in parenthesis using `\citet{}` (as in "See Hinton et al. (2006) for more information."). Otherwise, the citation should be in parenthesis using `\citep{}` (as in "Deep learning shows promise to make progress towards AI (Bengio & LeCun, 2007).").

The corresponding references are to be listed in alphabetical order of authors, in the REFERENCES section. As to the format of the references themselves, any style is acceptable as long as it is used consistently.

4.2 FOOTNOTES

Indicate footnotes with a number¹ in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).²

4.3 FIGURES

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction; art work should not be hand-drawn. The figure number and caption always appear after the figure. Place one line space before the figure caption, and one line space after the figure. The figure caption is lower case (except for first word and proper nouns); figures are numbered consecutively.

Make sure the figure caption does not get separated from the figure. Leave sufficient space to avoid splitting the figure and figure caption.

¹Sample of the first footnote

²Sample of the second footnote

Table 1: Sample table title

| PART | DESCRIPTION |
|----------|-----------------------------------|
| Dendrite | Input terminal |
| Axon | Output terminal |
| Soma | Cell body (contains cell nucleus) |

You may use color figures. However, it is best for the figure captions and the paper body to make sense if the paper is printed either in black/white or in color.

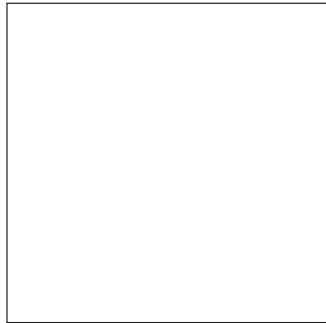


Figure 1: Sample figure caption.

4.4 TABLES

All tables must be centered, neat, clean and legible. Do not use hand-drawn tables. The table number and title always appear before the table. See Table 1.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

5 DEFAULT NOTATION

In an attempt to encourage standardized notation, we have included the notation file from the textbook, *Deep Learning* Goodfellow et al. (2016) available at https://github.com/goodfeli/dlbook_notation/. Use of this style is not required and can be disabled by commenting out `math_commands.tex`.

Numbers and Arrays

| | | |
|-----|---|---|
| 162 | a | A scalar (integer or real) |
| 163 | \mathbf{a} | A vector |
| 164 | \mathbf{A} | A matrix |
| 165 | \mathbf{A} | A tensor |
| 166 | \mathbf{I}_n | Identity matrix with n rows and n columns |
| 167 | \mathbf{I} | Identity matrix with dimensionality implied by context |
| 168 | $\mathbf{e}^{(i)}$ | Standard basis vector $[0, \dots, 0, 1, 0, \dots, 0]$ with a 1 at position i |
| 169 | $\text{diag}(\mathbf{a})$ | A square, diagonal matrix with diagonal entries given by \mathbf{a} |
| 170 | \mathbf{a} | A scalar random variable |
| 171 | \mathbf{a} | A vector-valued random variable |
| 172 | \mathbf{A} | A matrix-valued random variable |
| 173 | Sets and Graphs | |
| 174 | \mathbb{A} | A set |
| 175 | \mathbb{R} | The set of real numbers |
| 176 | $\{0, 1\}$ | The set containing 0 and 1 |
| 177 | $\{0, 1, \dots, n\}$ | The set of all integers between 0 and n |
| 178 | $[a, b]$ | The real interval including a and b |
| 179 | $(a, b]$ | The real interval excluding a but including b |
| 180 | $\mathbb{A} \setminus \mathbb{B}$ | Set subtraction, i.e., the set containing the elements of \mathbb{A} that are not in \mathbb{B} |
| 181 | \mathcal{G} | A graph |
| 182 | $\text{Pa}_{\mathcal{G}}(\mathbf{x}_i)$ | The parents of \mathbf{x}_i in \mathcal{G} |
| 183 | Indexing | |
| 184 | a_i | Element i of vector \mathbf{a} , with indexing starting at 1 |
| 185 | \mathbf{a}_{-i} | All elements of vector \mathbf{a} except for element i |
| 186 | $\mathbf{A}_{i,j}$ | Element i, j of matrix \mathbf{A} |
| 187 | $\mathbf{A}_{i,:}$ | Row i of matrix \mathbf{A} |
| 188 | $\mathbf{A}_{:,i}$ | Column i of matrix \mathbf{A} |
| 189 | $\mathbf{A}_{i,j,k}$ | Element (i, j, k) of a 3-D tensor \mathbf{A} |
| 190 | $\mathbf{A}_{:,:,i}$ | 2-D slice of a 3-D tensor |
| 191 | \mathbf{a}_i | Element i of the random vector \mathbf{a} |

Calculus

| | | |
|-----|--|---|
| 216 | $\frac{dy}{dx}$ | Derivative of y with respect to x |
| 217 | $\frac{\partial y}{\partial x}$ | Partial derivative of y with respect to x |
| 218 | $\nabla_{\mathbf{x}} y$ | Gradient of y with respect to \mathbf{x} |
| 219 | $\nabla_{\mathbf{X}} y$ | Matrix derivatives of y with respect to \mathbf{X} |
| 220 | $\nabla_{\mathbf{x}}^2 f$ | Tensor containing derivatives of f with respect to \mathbf{x} |
| 221 | $\frac{\partial f}{\partial \mathbf{x}}$ | Jacobian matrix $\mathbf{J} \in \mathbb{R}^{m \times n}$ of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ |
| 222 | $\nabla_{\mathbf{x}}^2 f(\mathbf{x})$ or $\mathbf{H}(f)(\mathbf{x})$ | The Hessian matrix of f at input point \mathbf{x} |
| 223 | $\int f(\mathbf{x}) d\mathbf{x}$ | Definite integral over the entire domain of \mathbf{x} |
| 224 | $\int_{\mathbb{S}} f(\mathbf{x}) d\mathbf{x}$ | Definite integral with respect to \mathbf{x} over the set \mathbb{S} |
| 225 | | |
| 226 | | Probability and Information Theory |
| 227 | $P(\mathbf{a})$ | A probability distribution over a discrete variable |
| 228 | $p(\mathbf{a})$ | A probability distribution over a continuous variable, or over a variable whose type has not been specified |
| 229 | $\mathbf{a} \sim P$ | Random variable \mathbf{a} has distribution P |
| 230 | $\mathbb{E}_{\mathbf{x} \sim P}[f(\mathbf{x})]$ or $\mathbb{E}f(\mathbf{x})$ | Expectation of $f(\mathbf{x})$ with respect to $P(\mathbf{x})$ |
| 231 | $\text{Var}(f(\mathbf{x}))$ | Variance of $f(\mathbf{x})$ under $P(\mathbf{x})$ |
| 232 | $\text{Cov}(f(\mathbf{x}), g(\mathbf{x}))$ | Covariance of $f(\mathbf{x})$ and $g(\mathbf{x})$ under $P(\mathbf{x})$ |
| 233 | $H(\mathbf{x})$ | Shannon entropy of the random variable \mathbf{x} |
| 234 | $D_{\text{KL}}(P \ Q)$ | Kullback-Leibler divergence of P and Q |
| 235 | $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ | Gaussian distribution over \mathbf{x} with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ |
| 236 | | |
| 237 | | Functions |
| 238 | $f : \mathbb{A} \rightarrow \mathbb{B}$ | The function f with domain \mathbb{A} and range \mathbb{B} |
| 239 | $f \circ g$ | Composition of the functions f and g |
| 240 | $f(\mathbf{x}; \boldsymbol{\theta})$ | A function of \mathbf{x} parametrized by $\boldsymbol{\theta}$. (Sometimes we write $f(\mathbf{x})$ and omit the argument $\boldsymbol{\theta}$ to lighten notation) |
| 241 | $\log x$ | Natural logarithm of x |
| 242 | $\sigma(x)$ | Logistic sigmoid, $\frac{1}{1 + \exp(-x)}$ |
| 243 | $\zeta(x)$ | Softplus, $\log(1 + \exp(x))$ |
| 244 | $\ \mathbf{x}\ _p$ | L^p norm of \mathbf{x} |
| 245 | $\ \mathbf{x}\ $ | L^2 norm of \mathbf{x} |
| 246 | x^+ | Positive part of x , i.e., $\max(0, x)$ |
| 247 | $\mathbf{1}_{\text{condition}}$ | is 1 if the condition is true, 0 otherwise |
| 248 | | |
| 249 | | |
| 250 | | |
| 251 | | |
| 252 | | |
| 253 | | |
| 254 | | |
| 255 | | |
| 256 | | |
| 257 | | |
| 258 | | |
| 259 | | |
| 260 | | |
| 261 | | |
| 262 | | |
| 263 | | |
| 264 | | |
| 265 | | |
| 266 | | |
| 267 | | |
| 268 | | |
| 269 | | |

6 FINAL INSTRUCTIONS

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the REFERENCES section; see below). Please note that pages should be numbered.

7 PREPARING POSTSCRIPT OR PDF FILES

Please prepare PostScript or PDF files with paper size “US Letter”, and not, for example, “A4”. The `-t letter` option on `dvips` will produce US Letter files.

Consider directly generating PDF files using `pdflatex` (especially if you are a MiKTeX user). PDF figures must be substituted for EPS figures, however.

Otherwise, please generate your PostScript and PDF files with the following commands:

```
dvips mypaper.dvi -t letter -Ppdf -G0 -o mypaper.ps
ps2pdf mypaper.ps mypaper.pdf
```

7.1 MARGINS IN LATEX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the `graphicx` package. Always specify the figure width as a multiple of the line width as in the example below using `.eps` graphics

```
\usepackage[dvips]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.eps}
```

or

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

for `.pdf` graphics. See section 4.4 in the `graphics` bundle documentation (<http://www.ctan.org/tex-archive/macros/latex/required/graphics/grfguide.ps>)

A number of width problems arise when LaTeX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command.

AUTHOR CONTRIBUTIONS

If you’d like to, you may include a section for author contributions as is done in many journals. This is optional and at the discretion of the authors.

ACKNOWLEDGMENTS

Use unnumbered third level headings for the acknowledgments. All acknowledgments, including those to funding agencies, go at the end of the paper.

REFERENCES

- Yoshua Bengio and Yann LeCun. Scaling learning algorithms towards AI. In *Large Scale Kernel Machines*. MIT Press, 2007.
- Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT Press, 2016.
- Geoffrey E. Hinton, Simon Osindero, and Yee Whye Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18:1527–1554, 2006.

A APPENDIX

You may include other additional sections here.