Assignment 1 Analysis
ML 7641

Ryan Wingate

RWingate8
GTID# 903453413

Implementation Strategy

For this assignment, I used the **scikit-learn** Python library, specifically, version 0.21.2. Additional dependencies include Jupyter Notebook (4.5.0) for convenient iteration and inline documentation of code, pandas (0.25.1) for data manipulation and cleaning, and Matplotlib (3.1.1) for charts and visualization.

The package manager I used was Anaconda (4.7.11), and the Python version was 3.7.4. For additional details on replicating the environment used for this assignment, reference the README included with this report.

Specific Algorithm Implementations

For each of the algorithms outlined below, I used the appropriate scikit-learn implementation.

Decision Tree

DecisionTreeClassifier, imported via:
- `from sklearn import tree`

Neural Network

MLPClassifier, imported via:
- `from sklearn.neural_network import MLPClassifier`

Boosting

AdaBoostClassifier, imported via
- `from sklearn.ensemble import AdaBoostClassifier`

Support Vector Machine

NuSVC, imported via
- `from sklearn.svm import NuSVC`

K-Nearest Neighbors

KNeighborsClassifier, imported via
- `from sklearn.neighbors import KNeighborsClassifier`

Classification Problem Description

The problems I selected for this assignment are two famous datasets, both of which are designed for binary classification tasks. They have a similar number of features but differ in the number of available samples by a factor of roughly 50. This was intentional, so that I could experience the differences in training time between the various algorithms when the sample count varies widely.

Exact dataset preprocessing steps are available in Jupyter notebooks I provided in the submitted GitHub repository. A high-level overview of the data-preprocessing steps I completed:

- Discard irrelevant features,
- Fill nan values,
- Replace nominal and ordinal categorical data with numeric representations,
- Create dummy variables for all nominal categorical data with 3 or more possible categories (dropping one column, to prevent multicollinearity), and
- Normalize continuous data.

Following processing, for both datasets, I randomly extracted 20% of the data using scikit-learn's train_test_split function. I used the remaining 80% as training data. I used 20% of the training data as a cross-validation set.

Titanic Dataset

The first dataset is the well-known Titanic Survival dataset, downloaded from https://www.kaggle.com/c/titanic. I consider the dataset interesting for a few reasons:

- Ability to successfully learn a variety of ML algorithms, despite its small size,
- Accuracy values empirically shown to be in the high 70s to mid 80s, which falls into the "interesting" range for most datasets,
- Subject matter provides interesting social commentary,
- Possibility of deriving interesting features from raw data (Deck from Cabin, for example),
- Widely considered by the data science community as being "interesting,"

- Personal desire to work with this dataset as a data science rite of passage.

The raw Titanic dataset consists of 1309 samples with 12 features.

Table 1: Raw Titanic dataset features and types.

Feature	Data Type
PassengerId	Integer
Pclass	Category (3 value)
Name	String
Sex	Category (2 value)
Age	Float
SibSp	Integer
Par/Ch	Integer
Ticket	String
Fare	Float
Cabin	String
Embarked	Category (3 value)
Survived	Category (2 value)

I discarded the **Ticket**, **Name**, and **PassengerID** features as non-useful data, and used **Cabin** ("C123") to derive an 8-value classification **Deck** ("C"), where possible, then I discarded **Cabin** as well. **Sibs** and **Par/Ch** were the count of siblings and count of parents/children also on the boat, respectively. **Sibs**, **Par/Ch**, and **Pclass** were interpreted as Ordinal categories and left as integers in the data. **Survived**, **Pclass**, **Sex**, **Embarked**, and **Deck** were interpreted as Nominal categories and processed into dummy variables. **Age** and **Fare** were normalized.

Table 2: Processed Titanic dataset features and types.

Feature	Processed Data Type
Pclass	Ordinal Category
Sex	Nominal Category (2)
Age	Float
SibSp	Ordinal Category
Parch	Ordinal Category
Fare	Float
Deck	Nominal Category (8)
Embarked	Nominal Category (3)
Survived	Nominal Category (2)

Following processing, the dataset consisted of 1309 rows and 17 columns.

The task for the Titanic dataset is to predict whether or not a given passenger was likely to survive the Titanic disaster.

Adult Dataset

The second dataset is the Adult dataset, downloaded from https://archive.ics.uci.edu/ml/datasets/Adult. I consider the dataset interesting for a few reasons:

- Relatively large size is an interesting point of comparison with the small Titanic dataset,
- Accuracy values empirically shown to be in the high 70s to mid 80s, which falls into the "interesting" range for most datasets,
- Personal interest in the factors driving compensation,
- Availability of another, larger but very similar dataset ("Census," 200K+ samples) with similar information and actual compensation values, should I need more, or more precise, data, and
- Widely considered by the data science community as being "interesting," as evidenced by 1.6 million web hits since 2007, making it the second most popular dataset on the UCI ML dataset repository.

The raw Adult dataset consists of 46012 samples and 14 features.

Table 3: Raw Adult dataset features and types.

Feature	Data Type
age	Integer
employment-type	Category (8)
fnlwgt	Float
education	Category (15)
education-num	Integer
marital-status	Category (7)
occupation	Category (14)
relationship	Category (6)
race	Category (5)
sex	Category (2)
capital-gain	Float
capital-loss	Float
weekly-hours	Float
native-country	Category (41)
compensation	Category (2)

To reduce the overall dimensionality of the dataset, I also discard **fnlwgt**, **marital-status**, **relationship**, **capital-gain**, **capital-loss**, and **native-country**. I also discard **education-num** as being redundant with **education**. I exclude all samples that do not have a **workclass** or **occupation** assigned, or which have a **workclass** that has value "Without-pay," as all of these would be highly correlated with low income and therefore uninteresting

Due to the large number of available categories, I map:
• **education** to 6 categories from 15,
• **employment-type** to 3 categories from 8, and
• **occupation** to 13 from 14.

Employment-type, **occupation**, and **race** were interpreted as Nominal categories and processed into dummy variables. **Age** and **weekly-hours** were normalized.

Table 4: Processed Adult dataset features and types.

Feature	Processed Data Type
age	Integer
employment-type	Nominal Category (3)
education	Ordinal Category (6)
occupation	Nominal Category (13)
race	Nominal Category (5)
sex	Nominal Category (2)
weekly-hours	Float
compensation	Nominal Category (2)

Following processing, the dataset consisted of 46,012 rows and 24 columns.

The task for the Adult dataset is to predict whether a given person earns more than $50K.

Hyperparameter Tuning

Decision Tree

Titanic
Figure 1-1 shows the learning curve for a default decision tree trained on the Titanic dataset. The model exhibits the high variance characteristic of an overfit. This is expected as the default hyperparameters allow the decision tree to grow arbitrarily complex.
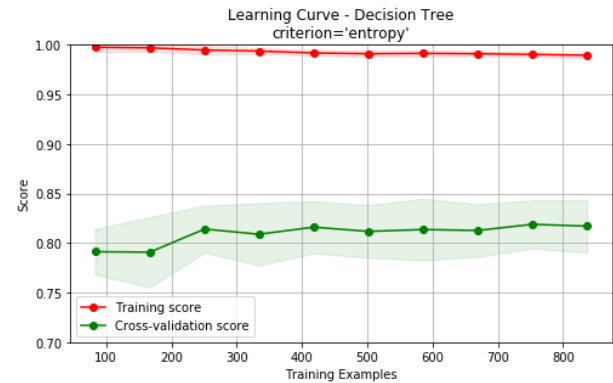

Fig 1.1: Learning curve for a decision tree with default hyperparameters, except for the split quality criterion.

Adult
The instructions herein are provided as a general guide for experienced Word users who would like to use that software to format their paper for an AAAI Press publication or report.

Error Rates

Titanic Dataset

The instructions herein are provided as a general guide for experienced Word users who would like to use that software to format their paper for an AAAI Press publication or report.

We assume that you can comply with the requirements as they are provided herein. If you are not an experienced Word user, please obtain assistance locally. AAAI cannot provide you with support. If the results you obtain are not in accordance with the specifications you received, you must correct your source file to achieve the correct result.

Results Analysis

Titanic Dataset

The instructions herein are provided as a general guide for experienced Word users who would like to use that software to format their paper for an AAAI Press publication or report.

Section Headings

The instructions herein are provided as a general guide for experienced Word users who would like to use that software to format their paper for an AAAI Press publication or report.
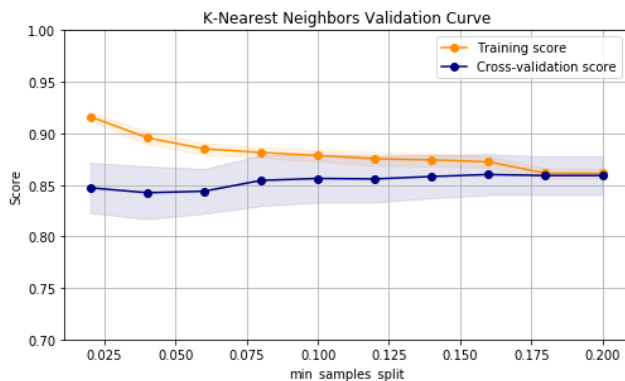
Section Headings

These instructions are generic. Consequently, they do not include specific dates, page charges, and so forth. Please consult your specific written conference instructions for details regarding your submission. Please review the entire document for specific instructions that might apply to your particular situation. All authors must comply with the following:

- You must use the latest AAAI Press Word template.
- Download the author kit.
- Complete, sign, and return by the deadline the AAAI copyright form (proceedings authors) or distribution license (technical report authors).
- Read and format your paper source and PDF according to the formatting instructions for authors.
- Submit your electronic files and abstract using the AAAI Press electronic submission form **on time.**
-

Email your copyright form, and any required page or formatting charges to AAAI Press so that they are received by the deadline.

- Check every page of your paper before submitting it.

K-Nearest Neighbors Validation Curve

- (This is the Bulleted List Style)

Copyright (This Is a Section Heading)

All papers submitted for publication by AAAI Press must be accompanied by a valid signed copyright form or, in the case of technical reports, by a valid signed permission to distribute form. There are no exceptions to this requirement. You must send us the original version of this form. However, to meet the deadline, you may fax (1-650-321-4457) or scan and e-mail the form (pubforms19@ aaai.org) to AAAI by the submission deadline, and then mail the original via postal mail, to the AAAI office. If you fail to send in a signed copyright or permission form, your paper cannot published. You will find PDF versions of the AAAI copyright and permission to distribute forms in the author kit. (This is an example of the Text style. Note that it is justified, and that hyphenation is turned on.)

(This is an example of the Text-Indent style. It is identical to the Text style, except that each paragraph is automatically indented.)

Formatting Requirements in Brief

We need source and PDF files that can be used in a variety of ways and can be output on a variety of devices. AAAI imposes some requirements on your source and PDF files that must be followed. Most of these requirements are based on our efforts to standardize conference manuscript properties and layout. These requirements are as follows, and all papers submitted to AAAI for publication must comply:

- **ALL FONTS MUST be embedded in the PDF file.** (Your paper cannot be corrected or published if the fonts are missing.)
- No type 3 fonts may be used (even in illustrations).
- All text in your paper must be black. Use of any colored type must be restricted to figures.
- Your title must follow mixed-case capitalization rules (that means all verbs (including short verbs like be, is, using,and go) nouns, adverbs, adjectives should be capitalized, including both words in hyphenated terms, while articles, conjunctions, and prepositions are lower case unless they directly follow a colon or long dash).
- Word documents must use the Times, Time Roman, or Times New Roman fonts. (Use of any other fonts in the text, title, heading, or references is not permitted).
- Fonts that require non-English language support (CID and Identity-H) must be converted to outlines or removed from the document (even if they are in a graphics file embedded in the document).
- Two-column format in AAAI style is required for all papers.
- The paper size for final submission must be US letter (8-1/2 x 11 inches). **No exceptions.**
- The source file must exactly match the PDF.
- The document margins must be as specified in the formatting instructions. Ensure that the specifications are for the whole document, not just the first section.
- The number of pages and the file size must be as specified for your event.
- No document may be password protected.
- Neither the PDFs nor the source may contain any embedded links or bookmarks. **URLS in your document must be black, and may not be underlined.**
- Your source and PDF must not have any page numbers, footers, or headers.
- Your PDF must be compatible with Acrobat 5 or higher.

If you do not follow the above requirements, it is likely that we will be unable to publish your paper.

What Files to Submit

You must submit the following items to ensure that your paper is published:

- A fully-compliant PDF file with complete PDF metadata.
- Your Word "docx" source file (required).

Your Word source will be reviewed and your PDF may be recompiled on our system. Name your source file with your last (family) name.

Do not send files that are not actually used in your paper. We don't want you to send us any files not needed for compiling your paper, including, for example, this instructions file, unused graphics files, additional material sent for the purpose of the paper review, and so forth.

Using Word to Format Your Paper

AAAI Press has provided a Word template that you can use to create your paper. You must be careful, however, not to change the page set-up of this document (print a PDF and use "shrink to fit" if you need to print it on A4 paper), and you will encounter problems if you use Identity-H or CID fonts. If your paper contains many in-line equations, and a significant amount of display mathematics, you may achieve better results using LaTeX, although the learning curve for this program is significantly higher. AAAI does not offer support in the use of Word or LaTeX.

You are required to use the above custom style sheets embedded in this document without alteration **Do not** use the built-in Word styles such as Header, Normal, and Heading 1.

Style Sheets (This is an example of a Subsection Heading)

The Word template provided by AAAI Press supplies custom styles for the most common elements in conference papers. **Use of the following style sheets is required:**

- Abstract Head
- Abstract Text
- Affiliation and Address
- Author Name
- Bulleted List
- Extract
- Figure Caption (must also be used for Tables)
- Footnote Text
- Section Heading
- Subsection Heading
- Text
- Text-indent
- Paper Title

Inserting Document Metadata with Word

PDF files contain document summary information that enables us to create an Acrobat index (pdx) file, and also allows search engines to locate and present your paper more accurately. *Inserting metadata is a requirement of submission.*

To insert metadata, open the document properties window. Type the title exactly as it appears on the paper (minus all formatting). Input the author names in the order in which they appear on the paper (minus all accents), separating each author by a comma.

Important! Do not include any nonascii characters (including accented characters) in the metadata, even if the names have accent marks. The data in the metadata must be completely plain ascii. It may not include any slashes, accents, linebreaks, or unicode characters. You may also include keywords in the Keywords field. If you know the full title of the proceedings, include it in the subject line. Leave any additional metadata fields blank.

Word Copyright Notice

The copyright notice has been added to the Word template using an invisible, unnumbered footnote, appended to the example of a first-level heading. Don't delete it! Instead, carefully insert your own heading at the left of the existing type, then remove the old type, being careful not to remove the invisible footnote marker at the end of the line. Be sure the copyright date is correct. *(If you disable this footnote and transfer of copyright is required, your paper will not be included in the proceedings or digital library.)*

Size, Margins, and Column Width

Papers must be formatted to print in two-column format on 8.5 x 11 inch US letter-sized paper. The margins must be exactly as follows:

- Top margin: .75 inches
- Left margin: .75 inches
- Right margin: .75 inches
- Bottom margin: 1.25 inches

The Word template is automatically set with these margins. Be certain, however, that the document margins apply to the *entire* document, not just the first section.

Column Width and Margins (This is a Subsubsection Heading Example)

To ensure maximum readability, your paper must include two columns. Each column should be 3.3 inches wide (slightly more than 3.25 inches), with a .375 inch (.952 cm) gutter of white space between the two columns. The Word template will automatically create these columns for you.

Nothing may intrude into the outside margins — including boxes around figures. Documents and tables may span two columns, but single-column figures, tables, and equations must remain completely within the column width.

Overlength Papers

If your paper is too long, turn on hyphenation, which will reduce the space used by many lines. Next, shrink the size of your graphics. **No alterations to page layout are allowed.** Do not use formatting tricks to make your paper a certain length. Cut text or remove figures instead. If your conference allows it, pay for an extra page.

Type Font and Size

Your paper must be formatted using the style sheets embedded in the Word Template. They will automatically format your paper in 10 point Times, Times Roman or Times New Roman. We will not accept papers formatted using other fonts (except that Cambria may be used for mathematics only.) This document is formatted in 10 point Times New Roman.

Line spacing for text must be exactly 12 point. The style file for "Text" sets the line spacing at 12 point. You may not alter line spacing or point size.

Do not use type 3 fonts for any portion of your paper, including graphics. AAAI will not accept electronic files containing obsolete type 3 fonts. Files containing such fonts (even in graphics) will be rejected.

If you are unsure if your paper contains type 3 fonts, view the PDF in Acrobat Reader. The Properties/Fonts window will display the font name, font type, and encoding properties of all the fonts in the document. If you are unsure if your graphics contain type 3 fonts (and they are PostScript or encapsulated PostScript documents), create PDF versions of them, and consult the properties window in Acrobat Reader.

Start all pages (except the first) directly under the top margin. (See the next section for instructions on formatting the title page.) The Text-Indent style will automatically insert the proper indent when beginning a new paragraph. It should be used unless the paragraph begins directly below a heading or subheading. **Do not skip a line between paragraphs.**

Title and Authors

Your title must appear in mixed case (nouns, pronouns, and verbs are capitalized) near the top of the first page, centered over both columns. The Paper Title style will automatically apply the proper font size and leading to your title. There should be two carriage returns above the title. The Word template already includes this space.

Author's names should appear centered below the title of the paper in boldface. You must use the Author Name style, which will automatically apply the proper size, font, and spacing. If space allows, you may use tabs to separate them, and insert affiliations below the names. If you have many authors, do not use tabs; simply separate the authors with commas. You may then use superscript numbers as a key to their affiliations, which should appear in the affiliation lines below the author list.

Affiliations and contact information must appear centered below the author names block. The Affiliation and Address style will automatically apply the correct formatting.

Do not use a table to provide this information. Please also note that no part of the title, author, or affiliation block may intrude into the margin.

You should begin the two-column format when you come to the abstract.

Credits

Any credits to a sponsoring agency should appear in the acknowledgments section, unless the agency requires different placement. If it is necessary to include this information on the front page, use a separate footnote. The AAAI copyright notice (required) must appear first in any list of footnotes.

Abstract

All conference papers must include an abstract. The abstract must be placed at the beginning of the first column, indented ten points from the left and right margins. The title "Abstract" should appear in ten-point bold type, centered above the body of the abstract. The Abstract Head style will automatically format this heading properly.

The abstract should be set in nine-point type with ten-point leading. The Abstract Head style will automatically apply these settings. This concise, one-paragraph summary should describe the general thesis and conclusion of your paper. A reader should be able to learn the purpose of the paper and the reason for its importance from the abstract. The abstract should be no more than two hundred words in length. (Authors who are submitting short one- or two-page extended extracts should provide a short abstract of only a sentence or so.) **Do not include references in your abstract!**

Page Numbers

Do not *ever* print any page numbers on your paper.

Text

The main body of the paper must be formatted in ten-point with twelve-point leading (line spacing). The Text style in this document is 10 point with 12-point line spacing. The first line of text after a heading should not be indented.

Subsequent lines of text within the same section or subsection should be indented 10 points. The 10 point tab is automatically set in the Word template. The Text-indent style will automatically apply the indent without having to use a tab.

URLs and Word

Microsoft Word's default setting is to underline URLs and display them in another color. This causes the URLS to turn light gray when printed, and often makes them undecipherable. **This feature MUST be turned off**. Please also turn off the automatic linking feature in Word. Failure to do so will result in a page reformatting fee.

Automatic Links

Do *not* use Microsoft Word's automatic section and reference linking.

Citations

Citations within the text should include the author's last name and year, for example (Newell 1980). Append lowercase letters to the year in cases of ambiguity. Multiple authors should be treated as follows: (Feigenbaum and Engelmore 1988) or (Ford, Hayes, and Glymour 1992). In the case of four or more authors, list only the first author, followed by et al. (Ford et al. 1997).

Extracts

Long quotations and extracts should be indented ten points from the left and right margins. The "Extract" style provides this type automatically:

> This is an example of an extract or quotation. Note the indent on both sides. Quotation marks are not necessary if you offset the text in a block like this, and properly identify and cite the quotation in the text.

Footnotes

Avoid footnotes as much as possible; they interrupt the reading of the text. When essential, they should be consecutively numbered throughout with superscript Arabic numbers. Footnotes should appear at the bottom of the page, separated from the text by a blank line space and a thin, half-point rule.

Headings and Sections

When necessary, headings should be used to separate major sections of your paper. Remember, you are writing a short paper, not a lengthy book! An overabundance of headings will tend to make your paper look more like an outline than a paper.

Section heads should be twelve-point Times New Roman bold type, mixed case centered. The Section Heading style will automatically apply these settings (including the extra required spacing). Subsection headings should be eleven-point Times New Roman bold type. The Subsection Heading style will automatically apply the proper settings automatically. Do not skip a line between paragraphs. Sub-subsection headings should be ten-point Times New Roman bold type, mixed case, with twelve-point leading, flush left, with three points of additional space preceding them and no additional points of leading following them. The Subsubsection heading style will apply the appropriate formatting automatically.

Section Numbers

The use of section numbers in AAAI Press papers is optional.

Section Headings

Sections should be arranged and headed as follows:

Acknowledgments. The acknowledgments section, if included, appears after the main body of text and is headed "Acknowledgments." This section includes acknowledgments of help from associates and colleagues, credits to sponsoring agencies, financial support, and permission to publish. Please acknowledge other contributors, grant support, and so forth, in this section. Do not put acknowledgments in a footnote on the first page. If your grant agency requires acknowledgment of the grant on page 1, limit the footnote to the required statement, and put the remaining acknowledgments at the back. Please try to limit acknowledgments to no more than three sentences.

Appendices. Any appendices follow the acknowledgments, if included, or after the main body of text if no acknowledgments appear.

References. The references section should be labeled "References" and should appear at the very end of the paper (don't end the paper with references, and then put a figure by itself on the last page). A sample list of references is given later on in these instructions. AAAI style is required for references. Poorly prepared or sloppy references reflect badly on the quality of your paper and your research. Please prepare complete and accurate citations.

Illustrations and Figures

Figures, drawings, tables, and photographs should be placed throughout the paper near the place where they are first discussed. Do not group them together at the end of the paper. If placed at the top or bottom of the paper, illustrations may run across both columns. Figures (including boxes and rules) must not intrude into the top, bottom, or side margin areas. Figures must be inserted using your page-formatting software. Number figures sequentially, for example, figure 1, and so on. There must be a minimum of 6 points between the figure and the figure caption, and an

additional 12 points beneath the caption. There must be 12 points between text and figures.

The illustration number and caption should appear under the illustration. Leave some space between the figure and the caption and surrounding type; .25 inches should suffice. Captions should be presented in nine-point Times New Roman italic. Labels, and other text in illustrations must be at least nine-point type. The Figure Caption style will apply the appropriate formatting automatically.

The First AI Magazine Cover.
(This is an example of the required figure caption style.
The illustration uses the required Figure stylesheet)

Low-Resolution Bitmaps

You may not use low-resolution (such as 72 dpi) screen dumps and GIF files—these files contain so few pixels that they are always blurry, and illegible when printed. If they are color, they will become an indecipherable mess when converted to black and white. **Gif files and low-resolution jpg or png files should never be used.** The resolution of screen dumps can be increased by reducing the print size of the original file while retaining the same number of pixels. You can also enlarge files by manipulating them in software such as PhotoShop. **Your figures should be a minimum of 300 dpi when incorporated into your document.**

Using Color

The archival version of your paper will be printed in black and white and grayscale. Consequently, because conversion to grayscale can cause undesirable effects (red changes to black, yellow can disappear, and so forth), we strongly suggest you avoid placing color figures in your document. If you do include color figures, you must be mindful of readers who may happen to have trouble differentiating some colors. **Your paper must be decipherable without using color for distinction.**

Drawings

We suggest you use computer-drawing software (such as Adobe Illustrator or, (if unavoidable), the drawing tools in Microsoft Word) to create your illustrations. Do not use Microsoft Publisher. These illustrations will look best if all line widths are uniform (half- to two-point in size), and you do not create labels over shaded areas. Shading should be 133 lines per inch if possible. Use Times New Roman or Helvetica for all figure call-outs. **Do not use hairline width lines** — be sure that the stroke width of all lines is at least .5 pt. Zero point lines will print on a laser printer, but will completely disappear on the high-resolution devices used by our printers.

Photographs and Images

Photographs and other images should be in grayscale (color photographs will not reproduce well; for example, red tones will reproduce as black, yellow may turn to white, and so forth) and set to a minimum of 300 dpi. Do not pre-screen images.

References

Reference citations in the text should appear in author-year format, for example (Smith 1975). References of the same year by the same author(s) should be distinguished by small letters following the year, for example (Smith 1977c) and alphabetized alphabetically by title. When referring to a paper in text in narrative form, use a narrative form. For example, say "In his paper, Michael Youngblood (2017) refers to ...." not (Youngblood 2017) refers to...."

All entries in the reference list must be cited in the text. In-text citations of four or more authors should be shortened to "first author et al." For example, the university technical report reference in the sample that follows would be cited as (Vattam et. al. 2013) in the text because it contains four authors.

Generally, references include the name of the author (surname first, followed by initials only for given names) and the date, followed by a period, then the title, presented in mixed case. For multiple authors, separate two names with a comma, and three or more authors with a semicolon. The place of publication (which is required for all book and proceedings publications) is followed by a colon, with the name of the publisher following. For journal articles and serial publications, provide the volume and issue numbers as well as the page numbers. DOIs are required for serial publications if they have been assigned. For conference papers, and book chapters, give inclusive page numbers. Provide the DOI if it is available. Do not use

shorthand abbreviations (such as AAAI-19) — spell out the full title of the publication.

If you are citing an ephemeral or general page of a website (such as, but not limited to, the landing page of a company or product), please do not include the citation in the reference list. Instead, incorporate the URL into an endnote (do not ever include the protocol prefix). All references must contain author, title, and date information.

To meet CrossRef requirements, all cited journal articles within a reference list must include a DOI if one has been assigned.

References should be listed alphabetically (by surname of the primary author or main entry) at the end of the article. Multiple references by the same author(s) should be listed in ascending chronological order with the earliest reference first (for example, Matthews 1979 precedes Matthews 1986). Information for each reference should be in the sequence illustrated by the following examples.

Dissertation or Thesis
(Note: Include department and university):

Clancey, W. J. 1979b. Transfer of Rule-Based Expertise through a Tutorial Dialogue. PhD dissertation, Department of Computer Science, Stanford University, Stanford, CA.

Forthcoming Book
Clancey, W. J. Forthcoming. *The Engineering of Qualitative Models.* Redwood City, CA: Addison-Wesley Publishing Company.

Preprint Server
Agrawal, A.; Batra, D.; and Parikh, D. 2016. Analyzing the Behavior of Visual Question Answering Models. arXiv preprint. arXiv:1606.07356v2 [cs.CL]. Ithaca, NY: Cornell University Library.

Published Book
Petroski, H. 1985. *To Engineer Is Human: The Role of Failure in Successful Design.* New York: St. Martin's Press.

Chapter in Published Book
Brown, J. S. 1977. Artificial Intelligence and Learning Strategies. In *Learning Strategies,* edited by J. O'Neil, 345–78. New York: Academic Press.

Forthcoming Journal Article
O'Connor, J. L. Forthcoming. Artificial Intelligence and Commonsense Reasoning. *AI Magazine* 44(3).

Published Journal or Magazine Article
Cox, M. T. 2007. Perpetual Self-Aware Cognitive Agents. *AI Magazine* 28(1): 32–45. doi.org/10.1609/aimag.v28i1.2027.

Paper Presented at Meeting
(Note: Use this format only if no published proceedings appeared):

Schoenfeld, A. H. 1981. Episodes and Executive Decisions in Mathematical Problem Solving. Paper presented at the 1981 AERA Annual Meeting. Boston, MA, September 24–30.

Zhou, S.; Suhr, A.; and Artzi, Y. 2017. Visual Reasoning with Natural Language. Paper presented at the AAAI 2017 Fall Symposium on Natural Communication for Human-Robot Collaboration. Arlington, VA, November 9–11.

Paper Presented at Meeting and Published in Proceedings
Lester, J.; Converse, S.; Kahler, S.; Barlow, T.; Stone, B.; and Bhogal, R. 1997. The Persona Effect: Affective Impact of Animated Pedagogical Agents. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems.New York: Association for Computing Machinery. doi.org/10.1145/258549. 258797.

Company Technical Report
Carbonell, J. R. 1970. Mixed-Initiative Man-Computer Instructional Dialogues, Technical Report QW-19871. Marina del Rey, CA: USC/Information Sciences Institute.

Scholarly Society Technical Report
Lin, F. 2007. Finitely-Verifiable Classes of Sentences. In *Logical Formalizations of Commonsense Reasoning: Papers from the 2007 AAAI Spring Symposium.* Technical Report SS-07-05. Palo Alto, CA: AAAI Press.

University Technical Report
Vattam, S.; Klenk, M.; Molineaux, M.; and Aha, D. W. 2013. Breadth of Approaches to Goal Reasoning: A Research Survey. In *Goal Reasoning: Papers from the ACS Workshop,* edited by D. W. Aha, M. T. Cox, and H. Muñoz-Avila. Technical Report CS-TR-5029. College Park, MD: University of Maryland, Department of Computer Science.

You are required to use the above AAAI reference format in your paper. If you fail to do so, your paper will be returned to you for reformatting (and a resubmission fee will apply). In addition, you must apply the buildt-in References style (which will automatically format your references in 9 point Time Roman with 10 point line spacing, and 3 additional points of space between each entry).

For the most up to date version and complete version of the AAAI reference style, please consult the *AI Magazine* Author Guidelines at

aaai.org/ojs/index.php/aimagazine/about/submissions #authorGuidelines.

Proofreading Your PDF

Please check all the pages of your PDF file. Is the page size A4? Are there any type 3, Identity-H, or CID fonts? Are all the fonts embedded? Are there any areas where equations or figures run into the margins? Did you include all your figures? Did you follow mixed-case capitalization rules for your title? Did you include a copyright notice? Do any of the pages scroll slowly (because the graphics draw slowly on the page)? Are URLs underlined and in color? You will need to fix these common errors before submitting your file. A more complete list of common Word formatting errors can be found in the Common Formatting Errors document, which is located in the 2019 AuthorKit.

Improperly Formatted Files

In the past, AAAI has corrected improperly formatted files submitted by the authors. Unfortunately, this has become an increasingly burdensome expense that we can no longer absorb (we are charged double for papers that require reformatting). Consequently, if your file is improperly formatted, it will probably be returned to you by the outside Production agency. If that happens, you will be required to fix your file and pay a resubmission fee.

Naming Your Electronic File

We require that you name your Word source file with the last name (family name) of the first listed author so that it can easily be differentiated from other submissions. Complete file-naming instructions will be provided to you in the submission instructions.

Submitting Your Electronic Files to AAAI

Submitting your files to AAAI will be provided to you with submission instructions for your conference. Please consult this document for details on how to submit your paper.

Inquiries

If you have any questions about the preparation or submission of your paper as instructed in this document, please contact AAAI Press at the address given below. If you have technical questions about implementation of the aaai style file, please contact an expert at your site. We do not provide technical support for Word or any other software package. To avoid problems, please keep your paper simple, and do not incorporate complicated macros and style files.

AAAI Press
2275 East Bayshore Road, Suite 160
Palo Alto, California 94303
Telephone: (650) 328-3123
E-mail: See the submission instructions
for your particular conference or event.

Acknowledgments

The preparation of the files that implement these instructions was supported by The Live Oak Press, LLC, and AAAI Press.

Thank you for reading these instructions carefully. We look forward to receiving your electronic files!