A Specific, Concise, and Descriptive Title

line 1: 1st Given Name Surname   
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCID  
  
line 1: 2nd Given Name Surname  
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCID  
  
line 1: 3rd Given Name Surname  
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

*Abstract*—In this case study we used the skills and methods we have learned in linear algebra and MATLAB to classify handwritten digits. Specifically, we utilize k-means clustering and outlier detection techniques to determine the numerical meaning of the handwritten digits based on their greyscale vector representations. We used MATLAB to build a method that takes an image of a handwritten digit and determines what number, from 0 to 9, is depicted.

# Introduction (*Heading 1*)

This project is borrows the classic K-Means clustering algorithm to approach sorting the MNIST hand-written digits dataset. As this project utilizes a myriad of linear algebraic techniques and topics learned in ESE 105, it enforces the learning done in class and allows a practical application of the subjects explored. Furthermore, this project could serve as a strong foundation for further exploration of machine learning techniques in the future, where a contrast of K-Means against other unsupervised learning or even supervising learning techniques to approach the problem could lead to varying accuracy in sorting. Overall, this project aims to label a list of vector-represented images of handwritten digits based on their numerical representation, from 0-9. It also aims to label any outliers, which the algorithm deems far from any centroid.

# Methods

Our overall approach to sorting the images comes in several stages.

## Training the K-Means Algorithm to find optimal centroids

First of all, we must use the given training dataset of 1500 images to train our k-means clustering algorithm. To briefly summarize k-means clustering, *k* centroids are first initially randomly assigned to random images in the space, since our images were defined with 784 pixels. Each image is then assigned to its closest centroid based on the nearest neighbor metric, by finding a centroid that minimizes the Euclidean distance, or the two norm, between the given image and centroid.



The centroid of each cluster is then redefined as the geometric mean of the locations of all the images assigned to that vector. All images are then reassigned to their closest centroid, and this process repeats until the algorithm eventually converges when the locations of centroids do not change from one iteration to the next. We have also artificially defined a maximum iteration, after which we will manually terminate the algorithm.

Before the algorithm may be executed, however, we must define the hyperparameter , or the number of clusters used in the algorithm. However, there is not necessarily an optimal method to solve for the optimal ­ value for any given dataset, and any visualizing technique through which we may visually estimate the number of clusters present in the data is rendered unapplicable due to the large number of dimensions present in the data. Instead, we have simply resorted to trial and error to find a range of possible values which yields a high accuracy in testing.

The choice of must be balanced

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

## Some Common Mistakes

Delete this subsection once you read it.

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

## Figures and Tables

Delete this subsection once you read it.

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

# Results and Discussion

Show the results that you achieved in your work and offer an interpretation of those results. Acknowledge any limitations of your work and avoid exaggerating the importance of the results.

# Conclusion

Summarize your key findings. Include important conclusions that can be drawn. Discuss benefits or shortcomings of your work and suggest future related project ideas you might like to explore in the future.

##### References

Provide citation information for all the previous publications referred to in your paper. Cite only those references that directly support your work.

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**