Using ML to help Predict Particle Collision

Halle Purdom1, Jobin Joseph1, Kebur Fantahun,Nibhrat Lohia2

1 Master of Science in Data Science, Southern Methodist University,  
Dallas, TX 75275 USA

hpurdom@smu.edu, jobinj@mail.smu.edu, kfantahun@mail.smu.edu

nlohia@mail.smu.edu

**Abstract.** In this paper we hope to leverage ability of data science to unlock secrets about how particle collisions work. There is a facility in Europe (CERN) that has built a collider underground which can produce collision of particles under high speeds. The problem is that there are many detectors on the collider which produce a great amount of data. In this paper we are trying to decipher what factors are important to find the particle trajectories and predict future collisions.

1 Introduction

1. What is the topic? Why should we care? Our topic is about particle collisions. Particle collisions can inform us about what the universe is made of and how it works. We could find new particles, look at the properties of resultant particle debris and hopefully discover new physics. ​
2. Why is this topic important to investigate?
3. What have the other experts discovered about the topic? There are many physicist around the world who have made discoveries using the particle colliders
4. What are the gaps? What is the problem?
5. How will this research solve this problem?
   1. This research aims to use \_\_\_\_\_\_\_\_\_ to solve \_\_\_\_\_\_\_\_\_\_.

2 Literature Review

This section is on the background information to help to reader understand why this research’s method will work.

**FOR THE OUTLINE ONLY – You do an annotated bibliography (1 paragraph summary of each article that will be used in the paper).**

2.1 Theme

2.2 Citations -READ ME

The list of references is headed “References” and is not assigned a number. The list should be set in small print and placed at the end of your contribution, in front of the appendix, if one exists. Please do not insert a pagebreak before the list of references if the page is not completely filled. An example is given at the end of this information sheet. For citations in the text please use square brackets and consecutive numbers: [1], [2], [3], etc. Use **APA format** in the reference section. You can choose to either have it alphabetical order or order of which it is shown in the paper.

**Hypothesis at the end of your literature Review**

3 Methods

1. Data
2. We are planning to get our data from CERN Open Data https://opendata.cern.ch/record/12303
3. Methods we plan on using MVA methods including Manova, PCA, LDA and Multidimensional scaling.

4 Results

1. What you hope to find in your research? Accept or reject the hypothesis

\*\*This Section is for statistical jargon and tables/Figures. Results are facts.

5 Discussion

\*\*\*Do not add New Results. This section is to apply and interpretate the results into lay terms.

\*\*\* Write questions you hope to answer in your research.

1. Interpretations: What do the results mean?
2. Implications: Why do the results matter? How should the reader apply these findings?
3. What stood out as interesting/unique/unexpected?
4. Limitations
   1. What challenges occurred during analysis?
5. Ethics
6. Future Research
   1. Are there areas of research where others can pick up and go deeper?

6 Conclusion

2 paragraphs max on the overall findings and summary of the research.

Acknowledgments. The heading should be treated as a 3rd level heading and should not be assigned a number.

References

1. Acciani, C., Fucilli, V., & Sardaro, R. (2011). Data mining in real estate appraisal: A model tree and multivariate adaptive regression spline approach1. Aestimum, (58), 27-45.
2. Carlucci, M., Grigoriadis, E., Venanzoni, G., & Salvati, L. (2018). Crisis-driven changes in construction patterns: Evidence from building permits in a mediterranean city. Housing Studies, 33(8), 1151-1174. doi:10.1080/02673037.2017.1421910
3. Davanloo Tajbakhsh, S., Castillo, E. d., & Aybat, N. S. (2015). On convex optimization methods for fitting spatial statistical models to large data sets ProQuest Dissertations Publishing.
4. Development of a forecasting model to predict the downturn and upturn of a real estate market in the inland empire (2013). Universal-Publishers.com.
5. Dheeriya, P. (2009). Modeling volatility in california real estate prices. IUP Journal of Applied Economics, 8(1), 26-38.
6. Dokko, Y., & Edelstein, R. H. (1992). Towards a real estate land use modeling paradigm. (special issue on office markets). Journal of the American Real Estate & Urban Economics Association, 20(2), 199.

Appendix

Use if needed for additional information