**Term Project Blog Report Guidelines**

These are typical ingredients o­­­­­f a blog-based project report. (depending on your project, some items may not apply).

* Header Picture
* Title
* Team members
* Abstract: 1-2 Paragraphs
* Introduction & Background (all items should be high level overviews)
  + Problem being addressed and why it’s important
  + Related work
  + Outline of approach and rationale (high level)
  + Contributions or novel characteristics of project
* Data Collection/Description
  + Relevant Characteristics
  + Source(s)
  + Methods of acquisition
* Data Pre-Processing & Exploration
  + Feature engineering/selection
  + Relevant Plots
* Learning/Modeling
  + Chosen models and why
  + Training methods (validation, parameter selection)
  + Other design choices
* Results
  + Key findings and evaluation
  + Comparisons from different approaches
  + Plots and figures
* Conclusion
  + Summarize everything above
  + Lessons learned
  + Future work - continuations or improvements
* References
* Relevant project links (i.e. Github, Bitbucket, etc…)

**General Tips:**

* In general, imagine your audience has a basic understanding of machine learning concepts, but is likely far from an expert. It should be an easier read than most academic research papers.
* Key code snippets can be helpful for people trying to understand how you implemented your project.
* Be sure to include images, gifs, or short videos. These will help liven up your post, make it more attractive to readers, and give you an opportunity to flex your data visualization skills.
* Plenty of good examples can be found on Medium on TowardsDataScience.
* Make sure your writing is fluid and grammatically correct.

**Evaluation Criteria include**:

* Clear description of project goals (and business relevance if applicable)
* Approach to pre-processing of data and feature extraction, the choice of data mining models used (and rationale for these choices)
* New theory/math (if applicable, most projects won’t have this component)
* Intelligent selection and tuning of models, addressing overfitting vs underfitting
* Novelty of approach/method/algorithm (if applicable)
* Presentation and evaluation of results
* Replicability of the results (is the description such that someone well versed in the art can obtain similar results on the same data?)
* Insights obtained from the effort
* Potential business impact or how the results can be “actioned upon” (if applicable)
* Appropriate/relevant reference list
* Quality of the writing (grammar and style)
* Visuals to aid in understanding of project elements

**Examples:**

* <https://towardsdatascience.com/automatic-speaker-recognition-using-transfer-learning-6fab63e34e74>
* <https://towardsdatascience.com/predict-the-number-of-likes-on-instagram-a7ec5c020203>
* <https://towardsdatascience.com/youtube-views-predictor-9ec573090acb>
* <https://towardsdatascience.com/a-data-science-for-good-machine-learning-project-walk-through-in-python-part-one-1977dd701dbc>
* <https://towardsdatascience.com/predicting-school-performance-with-census-income-data-ad3d8792ac97>

**Instructions for Submission of Blog + Supplementary Materials**

You need to submit a single document (Google doc, Word or pdf) via Canvas, one per team, that contains:

1. Title and Team
2. A hyperlink to your blog
   1. Optional: If there is something you would like to express as part of your report but not include it in the blog (e.g. your experiences and takeaways working in a cloud environment, which you prefer not to present in public), you can link to a separate document (max 1 page).
3. Hyperlink to code (+ associated documentation) repository