FLATIRON SCHOOL

Student Name: Mindy Jen
Instructor: Yish Lim
Date: 10/29/2020
Attempt Number: 1

Data Science Online Data Science Bootcamp
Capstone Final Project Review

Technical Notebook

Project Specifications	Metric for success	Developing		Accomplished		Exemplary (X-Factor)		Notes
README.md	Student has a clear README, highlighting important aspects of the project.	Student does not have a readme, or has a readme that is just a copy of the notebook.	☐	Student has a readme with a clear and well organized outline, conclusion and recommendation section. Visualizations are present.	☐	Student has a readme with a clear and well organized outline, conclusion and recommendation section. Visualizations are present. Language and markdowns lend themselves to succinctness.	☑	Really well-documented Readme with images and hyperlinks to additional resources.
Pick a novel interesting problem at the appropriate challenge level.	The chosen dataset was relevant to deep learning and was applied to business.	Business case not clearly articulated. Answered an obvious business question.	☐	Business case constructed clearly. And answered an obvious question, like clearly articulated the business stakeholder requirements that the project aims to accomplish. Business case constructed clearly and answered in notebook. Contained 1 or 2 business recommendations that are supported by analysis.	☐	Created original and meaningful work - Created a unique business case for the chosen dataset. Business case constructed clearly and notebook contains 3 or more business recommendations that are supported by analysis.	☑	Sound and thoughtful business recommendations.
Preprocess data	Import the data and preprocess the data that includes cleaning, scrubbing, handling missing values, etc.	Data not fully ready for later analysis. 100% correctly structured data. Handled missing values.	☐	Explored different methods.	☐	Handled especially tricky issues. Explored different methods with benchmarking.	☑	Used skimage and openCV to deal with image data
Describe data	Use EDA to create meaningful visualizations that describe your data. Plotting words to show cosine similarity, showing plots for class imbalance, etc.	No visualizations are present in the notebook	☐	1 or 2 visualizations are present in the notebook and visualizations are relevant to the project in a technical or business sense.	☐	3 or more visualizations are present in the notebook and visualizations are relevant to the project in a technical or business sense.	☑	Harder to come up with visualizations for a project dealing w images, but great job w the model loss/accuracy visualizations.
Fit models/Hypothesis testing	Fit at least one model. Summarize model impact and meaning.	Attempted basic model fitting (or forgot to model fit). Incorrect application. Misinterpreted results.	☐	Correctly fit a single model. Correctly interpreted model results. Summarized model meaning & impact.	☐	Compared multiple models. Fit models outside of class materials. Detailed numerical and visual analysis of models.	☑	

Present to technical audience	Present work done to a technical audience with code, insights, summary, future work, and even a live demo (for extra credit).	Unintelligible, hard to follow. Unclear. Incomplete.	☐	Engaging talk with insights & lessons. Explained code examples.	☑	Live demo! Ran code and changed parameter values.	☐	Live demo NA
Write quality code	Code is non-repetitive and uses OOP when necessary to avoid repetition. Custom methods/classes contain docstrings to help the reader understand what is happening. Variables have names that are relevant to what they represent.	Code is unorganized, lacks docstrings, variables are not named intentionally, and code repeats itself.	☐	Code lacks docstrings, but does not repeat itself and uses custom methods to do repetitive tasks. Code follows pep-8 standards.	☐	Code follows pep-8 standards, contains docstrings/comments, does not repeat itself and uses custom classes methods for tasks.	☑	Really organized notebook, great job with the .py files. Check out pep-8 standard docstrings to make it even more professional
Conclusion	Notebook contains a conclusion with business recommendations that are driven by analysis.	No conclusion present.	☐	Conclusion present but only states findings and contains 1 or 2 relevant business recommendations.	☐	Conclusion is present and contains at least 3 recommendations that are business relevant.	☑	
X - factor: Did something out of the box	Went above and beyond to research some additional topic, concept, Python package(s).	Routine project. Repeated analysis covered in class/sections of the module.	☐	Showed creativity.	☑	Ground breaking.	☐	

Non-Technical Presentation

Project Specifications	Metric for success	Developing		Accomplished		Exemplary (X-Factor)		Notes
Present to non-technical audience	Present work done to a non-technical (business focused) audience with problem statement, business value, methodology explained simply, business recommendations, summary, and future work.	Unintelligible, hard to follow. Unclear. Incomplete. Slides are too verbose, slide notes non existent.	☐	Engaging talk with insights & lessons. Explained methodology. Slides have images, less text, slide notes present on slide that mirror the script of the presenter. One slide for each of the following - Problem statement, business value, methodology, business recommendations (each recommendation on a separate slide), future work/next steps.	☐	Additional slides like findings, or use of engaging images, graphics, material showing expertise in communicating to business stakeholders.	☑	
Slide Quality	Slides are light on text, engaging and tell a story.	Slides are very text heavy or highly unorganized and all over the place.	☐	Slides are organized and tell a story, but contain too much text at times, especially when a visualization will suffice.	☐	Slides are organized, contain visualizations that relay information and slides tell a story.	☑	
Duration	Your presentation should be between 5 and 8 minutes.	Presentation is over 10 minutes or under 3 minutes.	☐	Presentation is over 8 minutes or under 5 minutes.	☐	Presentation is between 5 and 8 minutes.	☑	

Non Technical	Presentation contains great data science that is delivered using non technical language.	Presentation uses technical terms without succinct explanations more than 3 times.	☐	Presentation uses technical terms without succinct explanations once or twice.	☑	Presentation does not use technical terms or provides succinct explanations when using them.	☑	Model Architecture slide - for a non-technical audience perhaps have an alternative for people who are not familiar with neural network terminology
Test Results	Hypothesis test results are shown and made relevant to the business, driving the recommendations from the project.	No tests are shown or tests shown do not relate to business.	☑	Test results are shown and made clear to business case.	☑	Test results are shown, made relevant to business case and also highlight deeper insights into the business.	☐	Would like to see your accuracy results that you spoke of - even if it's just at 50%. Is that comparable to a baseline?
Visualizations	Slides contain visualizations that take the place of text and give the viewer insight.	Slides do not contain visualizations or the visualizations present are not relevant to the story.	☐	Slides contain visualizations that are relevant to the story but hard to interpret.	☑	Slides contain visualizations that are relevant and easy to understand.	☐	Only model architecture shown
Recommendations	A great presentation contains business recommendations and steps moving forward.	No recommendations are made	☐	At least 3 recommendations are made, but are not driven by data analysis or model.	☐	At least 3 recommendations are made and are driven by analysis and model.	☑	Recommendations show your knowledge on how neural networks can be used and improved. Shows that you've put a lot of thought into how your project can evolve.
Future Work	A data scientist will never have enough time to explore all aspects of dataset. If you had more time, what other aspects of the dataset would you explore?	No slide on Future work.	☐	Future work slide content not well defined and/or articulated.	☐	Future work clearly articulated, explored, and its potential business impact (s) described.	☑	
Thank You Slide	Thank your audience for their time, it's a great practice.	Thank You Slide is not present.	☐		☐	Thank You Slide is present. Appendix includes additional work.	☑	

Qualitative Assessment

1. Problem Statement - how well was it defined for this project

A very technology-relevant project with great future applications laid out.

2. Things you did well:

An extremely technical project that goes into a lot of material that our curriculum doesn't cover. Great job doing your own research on CNNs and LSTMs.

3. Things to work on/ consider :

Any of your next steps would make great extensions to your project. Perhaps consider creating a front end?

4. Action items:

None!