# Solving a Healthcare Problem for Improved Service Delivery

|  |  |
| --- | --- |
| **TASKS & UNDERSTANDING** | **TO DO / COMMENTS** |

|  |  |
| --- | --- |
| Introduction |  |
| What is your understanding of the problem? | Keywords [Revise according to literature]   * Intelligent staff acquisition * Intelligent resource allocation * Demand forecasting   For Washington State Health Ministry |
| ACTUAL DESCRIPTION OF PROBLEM   1. Understand the demographic make-up of each area and nearby areas of the facilities to determine demand for types of health services. 2. Predict future demand for types of services in each area and nearby areas of the facilities. This involves looking at population increase/decrease forecasts, changing demographics, future infrastructure/job investments or population migration. | |
| ANSWER PATTERN  [<literature>] shows that {<feature1>, <feature 2>, <feature 3>, …} are shown to be correlated in {<metric>} way for (<key problem>) but `<limitation in literature>`. [2 lines] /<your understanding>/ [1-2 lines] | |

|  |  |
| --- | --- |
| Introduction |  |
| Can you write the introduction and the problem statement in formal conference paper-like format? | Make sure your submission is in an actual conference paper format.  Best place: NIPS – Overleaf  Alternate: CVPR for MSW  Write your submission using the NIPS format in Overleaf. |

|  |  |
| --- | --- |
| Assumptions |  |
| What assumptions would you make to simplify the problem? | This is trying to assess your technical understanding, do some research, see how current work has dealt with this part. |
| List of technical assumptions justifying your use of a specific $<model>$  List of general assumptions | |

|  |  |
| --- | --- |
| Current population counts in each area |  |
| How would you acquire the data for maternity and elderly populations in each area? |  |
| List of available datasets & their statistical descriptions  Is the data enough?  Address legalities restricting access to data [state policies allow access?] | |

|  |  |
| --- | --- |
| Current population counts in each area |  |
| Would you take any steps to process and refine the data? | Discussing Preprocessing  Discuss missing values  ?Data augmentation: Train a generative adversarial model to generate instances of new data to expand the dataset? |

|  |  |
| --- | --- |
| Future population counts in each area |  |
| Share your prediction model details or how you  would predict future populations for maternity populations in each area. | Describe your model’s technical details |

|  |  |
| --- | --- |
| References |  |
| Cite any libraries, APIs, publications. We believe in standing on the shoulders of giants. Please reuse any existing research papers, source code, libraries but make sure to cite them. | Cite work that others like to cite. Don’t rely on publications that no one reads, don’t ground your work on papers that have no intended impact. Research for the sake of research is quite useless and usually results in papers that have little to no impact.  Use JabRef / Mendley |

|  |  |
| --- | --- |
| Code |  |
| Share your R or Python source code. | Show your mastery of version control.  Show that you are able to write well-structured code.  Show you can use best practices for working with teams.  Look at their Github, use their code conventions  Look at udacity’s course on how to write Readmes within Github.  Final Submission via GitHub  Readme is an abstract  Below it you should have 2 hyperlinks  1st hyperlink redirects to a subdirectory of the github repo which has all the code and data  2nd hyperlink redirects to the actual report |

WHAT ACTUALLY COUNTS

|  |  |
| --- | --- |
| 2.1 CURRENT POPULATION COUNTS [A] | Methodology for Given Dataset |
| Describe how you would use the population data sources to extract current population counts  of maternity people and elderly people in each area. | How to extract {< CURRENT POPULATION COUNTS >} from #<dataset(s)>#. |

|  |  |
| --- | --- |
| 2.1 CURRENT POPULATION COUNTS [B] | Demo for Given Dataset |
| Give a short outline or coded example demonstrating the concept of how you would extract this information from this data source. | Demo of 2.1[A] |

|  |  |
| --- | --- |
| 2.1 CURRENT POPULATION COUNTS [C] | Methodology and Demo for Another Dataset |
| Are there any other freely, publicly available data sources you would consider using and how  would you extract information from these sources? |  |

Create a section in report for 2.1 do a comparison for datasets for {< CURRENT POPULATION COUNTS >}

Methodology

Block diagrams with descriptions and justifications

Link to demo.

|  |  |
| --- | --- |
| In context of feature {< CURRENT POPULATION COUNTS >} | |
| Given Dataset | Another Dataset |
| Methodology 1 for extraction of feature | Mythology 2 for extraction of feature |
| Link to demo 1 | Link to demo 2 |

|  |  |
| --- | --- |
| 2.2 FUTURE POPULATION COUNTS [A] | Methodology |
| Describe how you would use the population data sources to predict future additional  population/increase or decrease in population counts of maternity people and elderly people in  each area. | Decide prediction model(s)  LR / MLP  Discuss limitations of each and justify how you would do **due diligence** before actually using the model for predictions.  **due diligence:**   1. Technical    1. MLP capabilities and limitations    2. LR Capabilities and limitations 2. Problem Understanding    1. Demonstrate that you actually understand the problem 3. Future (Show your vision)    1. If applicable, LSTMs might do better forecasts because… |

|  |  |
| --- | --- |
| 2.2 FUTURE POPULATION COUNTS [B] | Demo |
| Give a short outline or coded example demonstrating the concept of how you would predict this from the population data sources. |  |

|  |  |
| --- | --- |
| 2.2 FUTURE POPULATION COUNTS [C] | Limitations of used features {<…>} in Given dataset and |
| Demographics alone give a limited picture of possible future populations.  For example, if job investments are planned for some locations this could result in more new families moving to these locations. Are there any other freely, publicly available data sources you would consider using and how would you extract information from these sources? | Skilled immigration data  ? |

**Target**

• Understanding of the problem: Data Science Dojo consultants often work on cutting-edge data science, machine learning, optimization, and mathematical modelling problems. Your ability to understand problem is critical to your success at Data Science Dojo.

• Rigor: How detailed your research was and how hard you tried matters more than the actual results.

• Attention to detail: Make sure that your document is not missing any details. Typos are never fun.

• Demonstration of skills: We are not looking for a unicorn. We don’t expect you to be outstanding in all the areas of the problem. We do, however, want you to be good at a few of the skills. Your

* coding skills;
  + data extraction
  + manipulation
* investigate and research data sources and publications;
* ability to model problems
* ability to communicate complex concepts.

**Task Distribution**

Yellow is actual code, spend most of your time on this

Blue is documentation, all your code and your findings are here, this is an important deliverable, even if your code works if the documentation is poor you won’t get your point across.

Pink is the research part, you can’t be expected to know the literature really well, don’t dive too deep into the papers. Just skim through them. Specifically, abstract, conclusion, methodology – just keywords and diagrams, specific parts of results and discussions specifically which discuss contributions and limitations

1. Create Repository [6th June] and structure the paper format, grab/create the datasets
2. Explore 2.1 [7th June]
3. Document your findings from 2. In the cvpr research paper format [8th June]
4. Explore the introduction and assumptions [8th June]
5. Document your findings in 4 [8th June]
6. Explore 2.2 [9th June]
7. Document your findings in 6. [9th June]
8. Polish your work [10th June] Have everything reviewed [10th June]
9. Submit [11th June]