**Minutes   
Machine Learning Model Development  
Meeting No. 8**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Location** | Swinburne Hawthorn Campus  Late Lab | | | | |
| **Date & Time** | Friday, September 13, 2024 4:00pm | | | | |
|  |  | | | | |
| **Contributors** | Henry Richardson Matthew CrossSeth Kalantzis | | | *Meeting Facilitator Secretary* | |
|  | | | | | |
| **Action Items from Previous Meeting** | | | | | |
| **Item** | | **Discussion Points** | | | |
| **Find three potential datasets for the team to review.**  ***Owner:*** *All*  ***Status:*** *Completed* | | * Team completed a thorough search for appropriate housing data. * Matthew created spreadsheet with features in all datasets he explored. The only viable dataset only contained information from 2016-2018. * Henry explored datasets outside of Australia. Identified a few American housing datasets which had potential. Team exploration of these datasets found that they either didn’t contain specific features of the houses (E.g., number of bedrooms), contained aggregated data (E.g., houses sold per year), or were targeted towards rental property availabilities and didn’t contain sales prices. | | | |
|  | | | | | |
| **New Items** | | | | | |
| **Item** | | **Discussion Points** | | | |
| **Should we pivot to a different topic area?**  ***Owner:*** *All*  ***Deadline:*** *n/a* | | * Team had exhausted all potential sources for appropriate housing data. * Discussion held around how we could make current data work. Ideas floated to pivot to different target for analysis (E.g., number of available houses instead of sales price). Decided that the sourced datasets could not offer a sufficiently detailed topic for analysis. * Discussion around which topic would be better. The team spent the next hour exploring datasets for each of the remaining topics, and was most impressed by the data available for *Air Quality and Health.* * Specifically, Seth identified two very strong sources of information for NSWs Health and Air Quality. * Explored both dataset sources in detail and felt satisfied that there a substantial amount of relevant and timely features to analyse. * Locked in new topic. | | | |
| **Search for additional datasets.**  ***Owner:*** *Henry*  ***Deadline:*** *Sept 17, 2024* | | * The team is to continue searching for relevant datasets over the next few days, to see if there is any more information we can integrate to deepen the analysis. | | | |
| **Data Collection.**  ***Owner:*** *Matthew*  ***Deadline:*** *Sept 17, 2024* | | * Relevant Air Quality Features to be selected, downloaded, and stored in shared repo. * Relevant Health Statistic datasets to be selected, downloaded, and stored in shared repo. | | | |
|  | | | | | |
| **Action Items** | | | | | |
| **Action Item** | | | **Owner** | | **Deadline** |
| * **Download Health and Air Quality Datasets** | | | *Matthew* | | *17 Sept, 2024* |
| * **Search for Additional Relevant Datasets** | | | *Henry & Seth* | | *17 Sept, 2024* |
|  | | | | | |
| **Next Meeting** | | | | | |
| **Location** | Swinburne Hawthorn Campus Room EN401 | | | | | |
| **Date & Time** | Tuesday, September 17, 2024 8:30am | | | | | |



|  |  |  |
| --- | --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Henry Richardson  *Meeting Facilitator* | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Matthew Cross  *Secretary* | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Seth Kalantzis  *Attendee* |