|  |  |  |  |
| --- | --- | --- | --- |
| BD Meeting Record | | | |
| Date | 2022.3.23 | **Time** | 19.00-21.00 |
| Venue | PMB 321 | **Recording person** | Chuyan Fu |
| Attendees | Shanglun Wu, Chuyuan Fu, FERDINAND Kevin（online）, Xi Yu | | |
| Absence（reasons） | No | | |
| Lateness（reasons） | No | | |
| Content | | | |
| 1. Discuss the required amount of data: should be less than a million, otherwise, it would take too long to train 2. Discuss the time period of data: we decided it should be between 2019-12-30 to recent days because the first public report of covid 19 was on 2019.12, and there has been a massive omicron breakout in recent weeks. 3. Discuss the aim of our project: Identify the relationship between public emotion on social networks and time. If possible, identify the relationship between covid vaccination rate and public emotion. 4. Discuss the allocation of workload: Shanglun will study the spark online lesson, Xi will analyse the usage of snownlp, Chuyan will crawl the data from Weibo. 5. Discussed some basic processes of data-preprocessing. | | | |
| Mission | **Principal** | **DDL** | **Notes** |
| Crawl data from Weibo | Download the data and create a CSV file | 2022.3.31 |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |