

# ML for python repository similarity

Printed from Asana

## Survey

- ☒ **Honglin Zhang: General method** due 4 Jun, 2023
  - Priority: Medium
  - Stage: In Progress
  - Notes: Together
- ☒ **Honglin Zhang: Embedding method** due 7 Jun, 2023
  - Priority: High
  - Stage: In Progress
  - Notes: Honglin
- ☒ **847805259@qq.com: LDA method** due 7 Jun, 2023
  - Priority: High
  - Stage: In Progress
  - Notes: Leyu
- ☒ **Honglin Zhang: Method for calculating similarity** due 8 Jun, 2023
  - Stage: In Progress
  - Priority: High
  - Notes: Together
- ☒ **Honglin Zhang: Context survey report** due 9 Jun, 2023
  - Priority: Medium
  - Stage: In Progress
  - Notes: Separately

## Project

- ☐ **Honglin Zhang: Dataset** due 13 Jun, 2023
  - Priority: High
  - Stage: Not Started
  - Notes: Together
- ☐ **Honglin Zhang: Collect dataset** due 13 Jun, 2023
  - Notes: Together
  - Stage: Not Started
  - Priority: High
- ☐ **847805259@qq.com: Pre-process dataset and analyse it** due 13 Jun, 2023
  - Notes: Together
  - Stage: Not Started
  - Priority: High
- ☐ **Honglin Zhang: Design** due 14 Jun, 2023
  - Notes: Together
  - Stage: Not Started
  - Priority: Medium
- ☐ **847805259@qq.com: Architecture** due 14 Jun, 2023
  - Notes: Together
  - Stage: Not Started
  - Priority: Medium
- ☐ **Honglin Zhang: Development** due 30 Jun, 2023
  - Notes: Separately
  - Stage: Not Started
  - Priority: Medium

- |   |                  |
|---|------------------|
| <input type="checkbox"/> <b>Honglin Zhang:</b> Learning embedding and transformer knowledge<br>Notes: Separately<br>Stage: In Progress<br>Priority: Medium      | due 30 Jun, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com:</b> Learning LDA knowledge<br>Priority: Medium<br>Stage: In Progress<br>Notes: Separately                         | due 30 Jun, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com:</b> Demo based on knowledge and dataset<br>Notes: Separately<br>Stage: Not Started<br>Priority: Medium            | due 30 Jun, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang:</b> Building metadata embedding model<br>Priority: Medium<br>Stage: Not Started<br>Notes: Honglin                    | due 20 Jul, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com:</b> Building LDA model<br>Priority: Medium<br>Stage: Not Started<br>Notes: Leyu                                   | due 20 Jul, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang:</b> Training and tuning models<br>Priority: Medium<br>Stage: Not Started<br>Notes: Together                          | due 20 Jul, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com:</b> Evaluation<br>Priority: Medium<br>Stage: Not Started<br>Notes: Together                                       | due 20 Jul, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang:</b> Comparison between embedding model and LDA model<br>Notes: Together<br>Priority: Medium<br>Stage: Not Started    | due 20 Jul, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com:</b> Interim demo<br>Priority: Medium<br>Stage: Not Started<br>Notes: Separately                                   | due 30 Jun, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com:</b> Extending and improving the projects<br>Priority: Low<br>Stage: Not Started<br>Notes: Separately and together | due 13 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang:</b> Final project<br>Priority: Low<br>Stage: Not Started<br>Notes: Separately  | due 15 Aug, 2023 |

## Dissertation

- |   |                  |
|---|------------------|
| <input type="checkbox"/> <b>Honglin Zhang:</b> Introduction<br>Priority: Low<br>Stage: Not Started<br>Notes: Separately | due 10 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang:</b> Literature review and survey   | due 30 Jun, 2023 |

Priority: High  
Stage: In Progress  
Notes: Separately

- |   |                  |
|---|------------------|
| <input type="checkbox"/> <b>Honglin Zhang: Designing of the project</b><br>Priority: Medium<br>Stage: Not Started<br>Notes: Together        | due 10 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang: Implementation of the project</b><br>Priority: Medium<br>Stage: Not Started<br>Notes: Separately | due 10 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang: Evaluation</b><br>Priority: Medium<br>Stage: Not Started<br>Notes: Together                      | due 10 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang: Conclusion</b><br>Priority: Low<br>Stage: Not Started<br>Notes: Together                         | due 10 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang: First draft</b><br>Priority: Low<br>Stage: Not Started<br>Notes: Separately                      | due 10 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang: Second edition</b><br>Priority: Low<br>Stage: Not Started<br>Notes: Separately                   | due 13 Aug, 2023 |
| <input type="checkbox"/> <b>Honglin Zhang: Final report</b><br>Priority: Low<br>Stage: Not Started<br>Notes: Separately                     | due 15 Aug, 2023 |
| <input type="checkbox"/> <b>847805259@qq.com: Demo introduction or poster</b><br>Priority: Low<br>Stage: Not Started<br>Notes: Together     | due 18 Aug, 2023 |