

04. LUI, Go Nam. Human age ranking from pairwise comparison data via HodgeRank.

*Summary of the report.*

In this final project, he tries to evaluate the capability of HodgeRank on the human age ranking by evaluating pairwise comparison data.

*Describe the strengths of the report.*

Different models include Uniform model, Bradley-Terry model, Thurstone-Mosteller model, and Angular transform model are explored.

*Describe the weaknesses of the report.*

More interpretation of the results can be included.

*Evaluation on quality of writing (1-5):*

3

'As we can see the trend is basically fits the real age.' Should be 'As we can see the trend basically fits the real age.'

*Evaluation on presentation (1-5):*

1

Presentation video can't be played.

*Evaluation on creativity (1-5):*

3

*Confidence on your assessment (1-3):*

2

## **01. LUI-GoNam\_poster**

### 1. Summary

The poster tried to evaluate the capability of HodgeRank on the human age ranking by evaluating pairwise comparison data. The author applied four different models and analyzed inconsistency.

### 2. Describe the strengths of the report.

The report compared different algorithms in the study of age ranking problem and also compared with the real age ranking

### 3. Describe the weaknesses of the report.

The report failed to give some algorithm explanation of the different models.

### 4. Evaluation on quality of writing (1-5): 5

The report is clearly written and well organized.

### 5. Evaluation on presentation (1-5): 4

### 6. Evaluation on creativity (1-5): 4

These study can give some inspiration for the ranking problems in our daily lifes.

### 7. Confidence on your assessment: 2

## Comment on paper 4

In paper 4, Human age ranking from pairwise comparison data via Hodge Rank, the author tried to evaluate the capability of Hodge Rank on the human age ranking by evaluating pairwise comparison data. They make the comparison between real age distribution and generated ranking, and evaluate the accuracy of people's judgement by investigating the RMSE.

The result is that four GLMs are tested and their inconsistencies are evaluated. HHRG is a suitable framework for managing the incomplete pairwise comparison data on the human age ranking. Based on the Error analysis and Inconsistency analysis, it is convinced that Thurstone-Mosteller model is the most appropriate GLM on human age ranking.

Strength: The strength of this paper is that the author uses reliable dataset includes 30 images from human age dataset.

Weakness: The weakness of this paper is the content is not too rich.

Evaluation on quality of writing: 4. The writing is clear and there is no obvious mistake. Pictures and charts are used in this paper.

Evaluation on presentation: 4. The paper is well organized and clear.

Evaluation on creativity: 3.

Confidence on your assessment: 3.

### **Group 4**

#### Summary of the report

Evaluate the capability of HodgeRank on the human age ranking by evaluating pairwise comparison data.

#### Strength

A relative sound process with prediction score, inconsistency analysis and error analysis.

#### Weakness

Only one model is tried.

Evaluation on quality of writing (1-5): 4

Maybe uniform the first person by changing 'we' to 'I', 'our' to 'my' since it's one person's work;

Evaluation on quality of presentation (1-5): 1

Link to video cannot be played.

Evaluation on quality of creativity (1-5): 3

Confidence on your assessment (1-3): 3

## . Human age ranking from pairwise comparison data via HodgeRank

- **Summary of this report:** This report evaluates the capability of HodgeRank on the human age ranking by evaluating pairwise comparison data. It is shown that the HodgeRank on random graph is a reliable and efficient framework for pairwise comparison data analysis on VQA.
- **Describe the strengths of the report:** The report shows a clear comparisons for different GLMs. The reasons supporting the results are well analyzed.
- **Describe the weaknesses of the report:** The conclusion in this report is not clearly organized.
- **Evaluation on presentation: 3**  
Sorry the video can't be played. Perhaps, it's deleted by Youtube :P  
The suggestion for the slide is that they should put some figures in the introduction part
- **Evaluation on Clarity and quality of writing (1-5): 4**  
(1) When we seek to construct a **complete (completely)** subjective
- **Evaluation on creativity (1-5): 4**  
The technical quality is great.
- **Overall ratings: 4**
- **Confidence on your assessment: 2**

Human age ranking from pairwise comparison data via HodgeRank

Summary:

Evaluate the capability of HodgeRank on the human age ranking by evaluating pairwise comparison data.

Strength of the project:

Very clear and concise explanation for the project. It is also good to analyse each image's absolute error to identify which images are relatively hard for the people to rank.

Weakness of the project:

The way the author standardize the ranking to get the z-score should be more clearly explained.

Evaluation on Clarity and quality of writing (1-5):	4
Evaluation on Technical Quality (1-5):	4
Overall rating:	4
Confidence on your assessment:	2

#### 04. LUI, Go Nam. Human age ranking from pairwise comparison data via HodgeRank

- **Summary:** This report tries to rank human age by pairwise comparison data with four generalized linear models: Uniform model, Bradley-Terry model, Thurstone-Mosteller model and Angular transform model. It analyzes model's inconsistency, too.
- **Strength:** The report performs inconsistency analysis and error analysis of different models instead of just showing the result.
- **Weakness:** It would be better to review the methods and give mathematical forms of the evaluation metric to help understand.
- **Evaluation:**

	Writing	Presentation	Creativity	Confidence
Score	3	3	3	2

04. LUI, Go Nam. Human age ranking from pairwise comparison data via HodgeRank.

**Summary:** The author evaluates the inconsistencies of four generalized linear models on human age ranking by HodgeRank.

**Strengths:** The report is well organized. The author study several models and make a comparison.

**Weakness:** it is not clear how the method is used in the experiment. It would be better if the author can describe the methodology in the report.

**Evaluation on quality of writing: 3**

**Evaluation on presentation: Cannot watch the video.**

**Evaluation on creativity: 3**

**Confidence on your assessment: 2**

#### 04.LUI-GoNam\_poster

Summary:

In this project, four GLMs are tested and their inconsistencies are evaluated. The work shows that HHRG is a suitable framework for managing the incomplete pairwise comparison data on the human age ranking.

This report gets a stable result by random forest regression. It is clearly written, and the figures look very good.

Evaluation on Clarity and quality of writing (1-5):	3
Evaluation on Technical Quality (1-5):	4
Overall rating:	4
Confidence on your assessment:	2

04. LUI, Go Nam. Human age ranking from pairwise comparison data via HodgeRank.

In this work, author evaluated human age ranking via Hodge Rank, global and z scores, inconsistency analysis, error analysis on four different models are evaluated. Through these experiments, author concluded that Thurstone-Mosteller model is the most appropriate model due to the best error performance.

Strengths: This work is a complete work on human age ranking and evaluations are conducted from three aspects.

Weakness: The idea of this work is common, and there is no further discussion on the methods when achieving different results in different property analysis.

Evaluation on quality of writing (3): The written and organization of the poster still have some space to be improved. Especially, author should make more discussion on your designed experiments, rather than just describe the result.

Evaluation on presentation (3): The presentation can also be improved, comparing to another group working on the same topic.

Evaluation on creativity (3): It is obvious that all the models achieved less than 50% accuracy, author should raise more problem on the dataset, for example, why it performed not well enough, or the method to improve such accuracy.

Confidence on your assessment(2)