Analysis Report: Age Assignment Model Comparison
Trust Stamp Assignment

Pablo Martínez-Agulló

September 2024

1 Introduction

Hanut, an online retailer, wants to enhance customer experiences by personalizing them based on age demographics. After exploring image-based age estimation solutions, they have shortlisted two models and are seeking our analysis to determine which model to adopt. Their consumer age profiles are segmented into specific brackets: 0-12, 13-15, 16-17, 18-24, 25-30, 31-40, 41-50, 51-60, 61-70, 71-80, and 81+.

The two shortlisted models are:

- **Model 1**
  - Predicts an age range, given as 'age_min' and 'age_max'.
  - Data available in 'data/model_1.csv', containing 6975 entries.
- **Model 2**
  - Provides a direct age prediction labeled as 'age'.
  - Data available in 'data/model_2.csv', containing 6969 entries.

The actual age data is stored in 'data/gt.csv' with 8644 entries. Both models' datasets contain some missing entries when compared to the ground truth data. After merging the three datasets, the comparison is conducted on 6957 complete entries, with missing data removed.

2 Data Exploration

On Figure 1 it can be already seen that as the real age goes higher, both models underestimate the age. The R^2 suggests a linear behavior, especially for Model 1. Note that on all plots, for the Model 1, the average between the 'age_min' and 'age_max'.

The horizontal alignment of points for Model 2 (blue) indicates that its predictions might be discretized to specific values instead of providing more granular continuous predictions. The low R^2 score of Model 2 may be caused by this.

When observing Figure 2:

Age Bin	**Model 1**	**Model 2**
0-12	Good	Good
13-15	Bad	Bad
16-17	Bad	Bad
18-24	Good	Good
25-30	Good	Good
31-40	Good	Good
41-50	Good	God
51-60	Good	Bad
61-70	Good	Bad
71-80	Bad	Bad
81+	Bad	Bad

Table 1: Comparison of Model 1 and Model 2 predictions across different age bins.

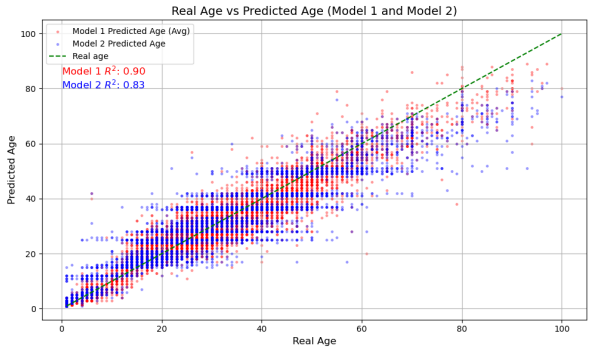

Figure 1: Scatter plot of predicted age for both models as a function of the real age.

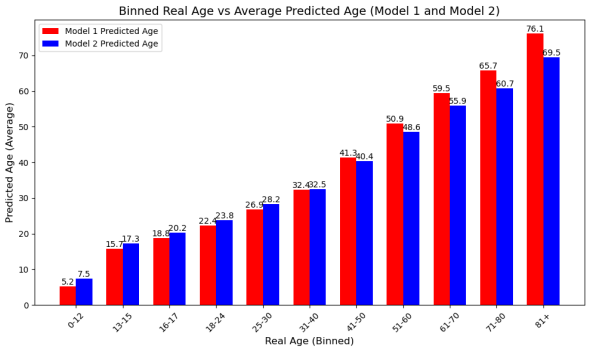

Figure 2: Bar plot comparing the average prediction of both models at each real-age bracket.

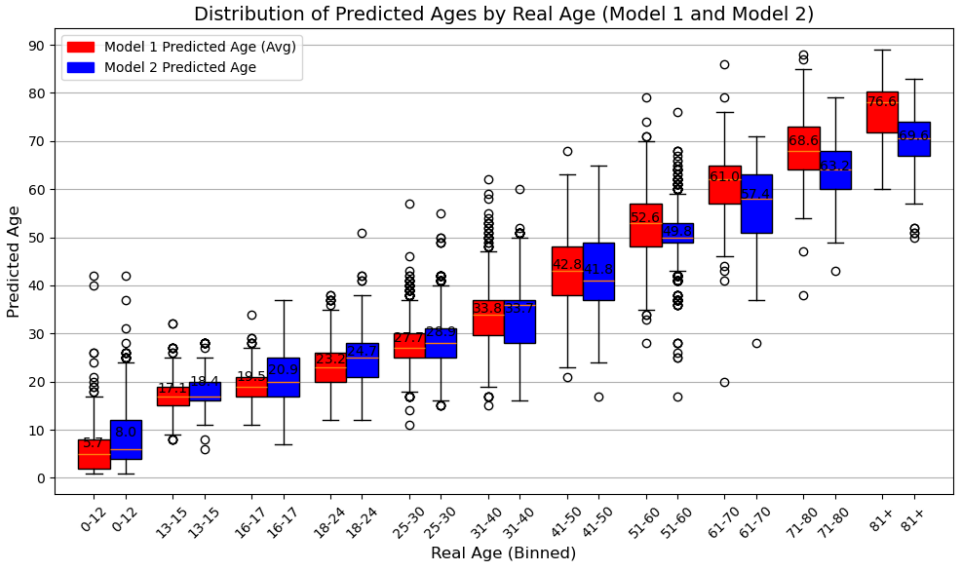

Figure 3: Caption describing the image.

3 Conclusion

(Here, you'd summarize findings and suggest which model is preferable based on performance metrics.)