



AI vs Marketing Experts – Who Analyzes Better?

Project Seminar 2024/2025

Introduction

AI is reshaping marketing analytics, enabling businesses to analyze demographics faster than ever. But can it match human expertise in recognizing age, gender, and ethnicity? While AI offers efficiency, it also risks misclassification and bias. This study explores where AI and humans align, differ, and what that means for real-world marketing decisions.

Project Objective

- Compare AI and human marketing experts in advertising analysis.
- Examine strengths, limitations, and biases in their classifications.
- Evaluate AI's effectiveness and its impact on marketing strategies.

Research Question

Can AI analyse advertising as effectively as human marketing experts, across demographic characteristics, do their perspectives reveal unique insights—or biases?

Data & Methodology

Variable	Description	Variable Values
Ethnicity	Specifies the ethnicity of the character.	1 = Caucasian/White; 2 = Black; 3 = Asian; 4 = South Asian; 5 = Mediterranean – Arabic or Middle Eastern; 6 = Mediterranean – Europe; 7 = Mediterranean – Europe; 8 = South/Latin American; 9 = Other
Gender	Specifies the gender of the character	1 = Female; 2 = Male; 3 = Child or Infant; 4 = non Human objects;
Age	Specifies the age gap of the character	1 = 18 or younger; 2 = 19-25; 3 = 26-35; 4 = 36-54; 5 = 55-70; 6 = 71+ / Not determined;

Quantitative Analysis

- Measuring agreement & disagreement rate
- Cohen's Kappa → Evaluates inter-rater reliability.
- Chi-Square test for independence → Checks if AI and human classifications are statistically related.
- Chi-Square test for bias & misclassification → Detects patterns in AI errors.

Qualitative Analysis

- Investigating divergences
- Model limitations → Identifying AI's classification challenges.
- Bias detection → Understanding systemic misclassification patterns.
- Human subjectivity → Evaluating inconsistencies in manual coding.

Results

AI model selection

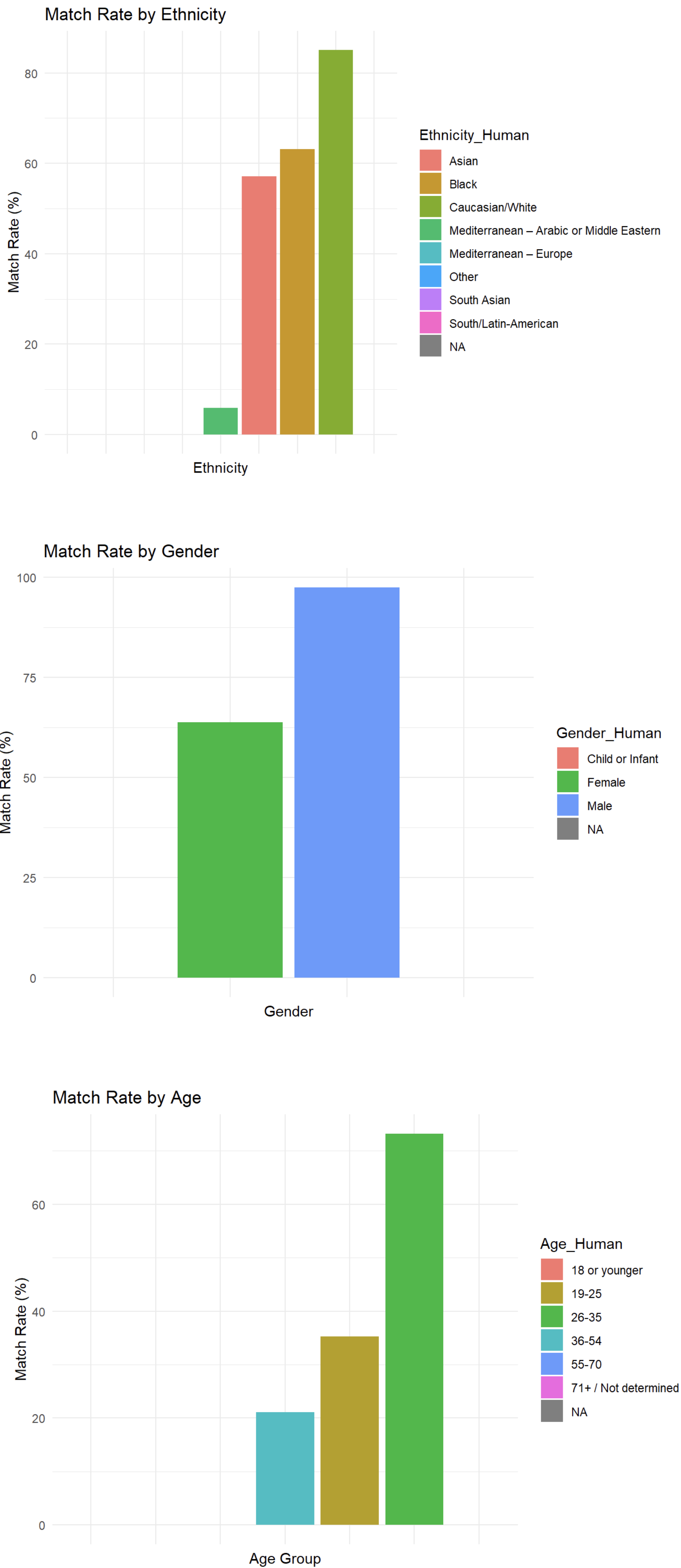
We finalized on DeepFace (MultiEmbed) model based on higher performance among 23 models.

Result comparison

	Ethnicity	Gender	Age
Agreement (%)	70.0	78.9	32.2
Disagreement (%)	30.0	21.1	67.8
(Weighted) Cohen's Kappa	0.214	0.575	0.263

Statistical test results:

- AI and human classifications differ significantly, especially in Ethnicity and Gender.
- Significant biases or misclassification from AI model. Ethnicity mismatches occur more frequently than Gender and Age in bias test.



Conclusion

Discrepancies:

- AI completely misclassifies children.
- AI tends to confuse certain ethnic categories, e.g. due to overlapping features.
- AI struggles with classifying extreme age groups

Alignments:

- Strong agreement on Caucasian/White ethnicity & Male gender.
- Mid-age groups (26-54) show high consistency between AI & human labels.

Key Patterns & Biases:

- AI doesn't always match human-labeled categories, causing misalignment.
- Bias detected – AI is more accurate for males, mid-aged, and dominant ethnic groups.
- Defaults to majority categories when uncertain.
- Cohen's Kappa is low, showing AI's agreement is often due to class imbalance.

R. Dennis, M. Nguyen, L. Jose Thottam, M. Egamkulov, V. Knurova, B. Naydenov.
Supervisor: Dr. Charlotte Köhler, Guidance: Dr. Anna Rößner