

## Group Assignment 2: Human-machine Interaction and Greater Zero-sum Gains [100 points]

Due: April 25th by 11:59 pm

### Instructions

In this assignment you will work in groups of 2-3 students max (they can be the same as the previous group assignment). You will be presented with an additional challenge: achieving complementary team performance. You will be designing an experiment to validate your solutions and hypotheses of what you think you are going to observe. Through a small-scale user study, you will collect some evidence to gain insights about your solution.

#### Deliverables:

1. A PDF containing all of your responses to the written questions shared in this notebook.

Upload the PDF in 1 to the main assignment link on Gradescope.

Only one submission per group is needed. Add the names of your team members:

[Venkata Datta Adithya Gadhamsetty, Mukund Iyengar, Gabriel De Romana]

#### Learning outcomes:

1. Formulating study designs to test your solutions.
2. Executing the study and interpreting the data acquired.

\**GitHub Link:* \* [https://github.com/dat-adi/xai\\_car\\_val](https://github.com/dat-adi/xai_car_val)

### Background

Make sure you read and understand these concepts before you start developing the assignment.

People team up with AI-powered systems to accomplish critical tasks as well as more mundane tasks. There is a premise stating that the expertise of the AI system complements the expertise of the human, which allows the partnership to accomplish even more than each actor involved alone. This is reflected in a better human-AI team performance or **complementary performance** because the interacting parties' aggregate gains and losses can be more than zero, i.e., there's a non-zero-sum situation.

In practice, the degree of complementary expertise can vary from no overlap of expertise to perfect complementarity. Besides, it is not always easy for a human user to know when an AI

perfect complementarity. Because, it is not always easy for a human user to know when a system may be an expert in some parts of a task but be a non-expert in others.

Despite observed improvements when humans collaborate with AI in certain tasks compared to humans working alone, there remains a gap with respect to the AI's performance in autonomous operation. The challenges mentioned above can potentially contribute to achieving such superior performance than both individual parts. Importantly, we would like to keep the human-in-the-loop for acceptance, accountability, and ethical and legal considerations.

Your goal: Design an AI assistance solution that will allow the human-AI team to achieve better task-related performance than each part operating separately.

Reference: [You Complete Me: Human-AI Teams and Complementary Expertise](#)

## ✓ Part 1 – Study overview

### Problem 1.1 [6 points]

Define an AI-based solution with explanations and specify the target end users. This problem should allow or **have the potential for greater zero-sum gains**. A task that is very easy for humans might not be suitable (it will be hard to beat the human alone). Because we are not following the initial stages of design thinking to identify real needs and define your solution again in this assignment, you can use any of the ideas from previous assignments.

Make sure to include

- The need being fulfilled by your solution
- What the model inputs, general mechanism, and outputs are
- How explainability or interpretability is incorporated

*Enter your answer here*

You may use the template below:

**Solution Overview:** An AI-powered tool that estimates second-hand car values by analyzing visible features and providing prototype-based explanations.

**Target End Users:** Car enthusiasts seeking to make informed second-hand car purchase decisions.

**Need Fulfillment:**

- Primary need addressed: Users need a fast, trustworthy estimate of used car value based on damages, features, and comparisons to similar cars.
- How this creates potential for complementary gains: AI can detect patterns (e.g., subtle

damage) humans might miss, while users can question the AI's valuation based on personal preferences.

#### Model:

- Inputs: Image of the car (uploaded by the user)
- Processing mechanism: The ProtoPNet model extracts visual parts, compares them to known prototypes (e.g., "wheel dent", "scratched door"), and predicts a value with part-wise explanations.
- Outputs: Predicted car resale value, List of detected issues with estimated price impact.

#### Explainability Approach:

- Technique(s): The technique would be a prototype explanation
- Implementation: The model highlights matched car parts to known prototypes and presents hypothetical value changes based on fixing specific damages.
- User interaction: Users upload the picture, view part-based breakdowns (prototypes), visualize which damages affected the price (explainability), and adjust scenarios to re-estimate the car value.

### ✓ Problem 1.2 [8 points]

Formulate the hypotheses you need to test in a user study that allow you to validate whether your solution achieves the main goal (complementary team performance). This requires specifying what you will be measuring (dependent variables) and the independent variables (experimental treatments or manipulations).

**Note:** balance the complexity and feasibility of running such a study. Think about what are the necessary conditions that would allow you to validate your solution.

*Enter your answer here*

For each hypothesis, complete the following:

**H1:** Prototype-based explanations effectively influence the user into correcting their estimate towards the ground truth.

- Rationale: Seeing specific parts that affected the price helps users understand and validate the model's reasoning, thereby enabling them to place more importance in them when re-estimating the price
- Relates to which aspect of complementary performance: Improved the user's trust in the prediction made by AI assistance (complementary expertise between human intuition and machine reasoning).

**H2:** Users supported by prototype-based explanations are more confident in their estimation

H2: Users supported by prototype based explanations are more confident in their estimation of the ground truth

- Rationale: The prototype-based explanation will increase the user's confidence in their prediction as compared to the case when they have no such AI assistance
- Relates to which aspect of complementary performance: Better joint decision performance stemming from increased confidence in the user's predictions (the human-AI team outperforms the human alone or AI alone).

...

Define the variables of the experiment in the tables below:

#### Independent Variables:

Variable Name	Type (Between/Within)	Levels/Conditions
User's estimated valuation	Within	Initial and post-explanation valuation estimation
User's confidence	Within	Confidence with and without an AI explanation

#### Dependent Variables:

Variable Name	Measures	Relates to Hypothesis
Weight of Advice	Magnitude of change in valuation with respect to the ground truth	H1
Change in confidence	Change in confidence pre and post explanation	H2

For the dependent variables, you don't need to specify how you will measure them here.

### ✓ Problem 1.3 [6 points]

Define your experimental task and how you will acquire the data to compute the dependent variables mentioned above.

*Enter your answer here*

Note: We recruit only experts as previous studies have shown that novices blindly trust AI systems, this makes it hard to measure influence.

**Experimental Task Description:** Initially, experts will be presented with an image of car and produce a blind valuation estimate after which they are requested for their confidence in their prediction. This leads into the second case where users are provided the same car, but with a prototype-based explanation to inform them of the model's prediction and the factors that constitute it. Post the user's second prediction, they are requested for their confidence in their now-informed prediction.

#### Data Collection Methods:

Dependent Variable	Collection Method	Calculation/Pi
Weight of advice	The user's initial and post-explanation estimate	The difference in user's estimation prior to an
Change in confidence	The user's confidence input pre and post explanation	The difference in the user's confidence of the

## ✓ Part 2 – Interface development

### Problem 2.1 [40 points]

Build your **React prototype** with the intent to gather meaningful feedback for evaluating your solution's effectiveness. We won't be asking you to build paper or Figma prototypes this time, but some planning of the interface and interactions is recommended.

Take into account that

- Your interface should support the recording of all the measurements relevant to validate your hypotheses.
- You may need additional pages to capture some metrics (e.g., through questionnaires).
- Data recording is a must here since you will recruit participants to validate your hypotheses.
- You don't need to host your web application. You can launch the web interface locally for the main study in the next stage.

The actual deliverable for this part will be a demo for the pitch session scheduled during the final exam slot of this course. More information about the format will be released soon.

As you complete this assignment, please prepare the demo.

*Provide a link to your code here. It can be a github repository or Google Drive folder*

### Bonus [10%]:

Implement the actual ML model that generates the outcomes of interest and connect it with your main interface. In this way, rather than mocking up the AI outcomes or behaviors, you can use your model to display the predictions and the corresponding explanations when needed.

Considerations:

- You may need access to certain datasets and annotations to train the model(s).
- Think about what information you need to access in the backend to be transferred to the frontend, where the user will have the interaction with your AI system.
- Real time processing is not required, but if your model relies on users' inputs to make some computations, then you may have to consider online processing.

If you decide to complete the bonus, please report the dataset, data splits, training setup, experiments, and model evaluation.

## ✓ Part 3 – Conducting a user study

### Problem 3.1 [5 points]

Before running your user study, we need to define the:

1. Experimental protocol of your study (execution steps)
2. Recruitment strategy (target population, inclusion/exclusion criteria, and participants' assignment).

As a guide, your responses should resemble the corresponding sections in the Methods sections of the papers we have discussed in class.

#### *One paragraph with your experimental protocol*

Participants will first complete an informed consent form and a short demographic survey. Each participant will then complete a series of car valuation tasks divided into two phases. In the first phase (baseline), participants will be shown an image of a second-hand car and asked to provide an initial estimated resale value, along with a confidence rating on a 1–10 scale. In the second phase (assisted), participants will be presented with the same image, now accompanied by a prototype-based AI explanation highlighting specific features of the car and a suggested AI valuation. Participants will then provide a revised valuation estimate and a second confidence rating.

#### *One paragraph with the participants information*

We will recruit 20 adult participants who have a strong personal interest in automobiles and experience purchasing cars, thereby possessing an informed understanding of the second-hand car market and valuation trends. Participants will not be professional car appraisers, but will be selected based on their demonstrated experience with the car market. Recruitment will be conducted through mutual contacts and online automotive communities. To ensure expertise, we screened participants by asking whether they have experience buying cars and whether they feel confident estimating car values based on market trends. Additionally, to avoid potential bias in confidence ratings, all participants were individuals not personally known to the researchers. No monetary compensation will be provided. We specifically chose to recruit experts and car enthusiasts, rather than novices, because prior research indicates that novice users tend to blindly trust AI systems, making it difficult to accurately measure the effects of AI explanations on user judgment.

## ✓ Problem 3.2 [35 points]

Collect data with at least 8 subjects. For this assignment, it is acceptable if the participants you recruit are not those specified as the target end users (explain if this is the case).

Note: normally, we would need IRB approvals for conducting human-subjects evaluation, even

when the risks are low because this type of studies involve intervention or testing (Neuropsychological/ cognitive/ psychosocial/ behavioral/ educational).

Do the statistical analyses and report your main findings. Did you find evidence for your hypotheses? briefly discuss your results.

*Enter your answer here*

**Sample Description:** Number of participants: N = 20

Relevant demographics: Participants were adults aged 21–26 years old. The sample included an even distribution of male and female participants (10 males, 10 females).

Recruitment source: Participants were recruited through mutual contacts to avoid bias in confidence rating.

Deviation from target population (if any): No major deviation. All participants met the inclusion criteria of having prior experience purchasing cars and an informed understanding of the car market. Participants were not professional appraisers but were car enthusiasts with practical experience, as intended.

**Statistical Analysis Plan:** Hypothesis H1 (Prototype explanations influence correction toward ground truth): Analysis: Paired-sample t-test comparing initial prediction errors (relative to ground truth) and final prediction errors. Dependent measure: Magnitude of prediction change.

Hypothesis H2 (Prototype explanations increase confidence): Analysis: Paired-sample t-test comparing confidence ratings before and after AI explanation. Dependent measure: Change in confidence score. Significance threshold:  $\alpha = 0.05$

Software used: Microsoft Excel

*Organize your results by sections considering the dependent variables or constructs measured. Include descriptive statistics and outcomes of the statistical tests. Graphs, tables, or visualizations can be included as well.*

You may use the template below:

**Results by Dependent Variable:**

✓ [Weight of Advice]

[Insert table of descriptive statistics] <https://docs.google.com/spreadsheets/d/1z3SYHVLFnjFpeMPGHsFhI3m4GHD0qcFRx7ZsVRsYC4/edit?usp=sharing>

Measure	Value
Mean Initial Prediction	6820
Mean Final Prediction	3170

Mean Prediction Change -3650

[Insert statistical test results] A paired-sample t-test revealed a statistically significant difference between participants' initial and final predictions,  $p = 4.435636 \times 10^{-11}$  ( $p < 0.001$ ).

[Optional: Insert visualization]

Key finding: [1-2 sentences highlighting the main result]

Participants' valuation predictions significantly shifted toward more accurate estimates after viewing the AI prototype-based explanations, supporting that the AI effectively influenced human judgment.

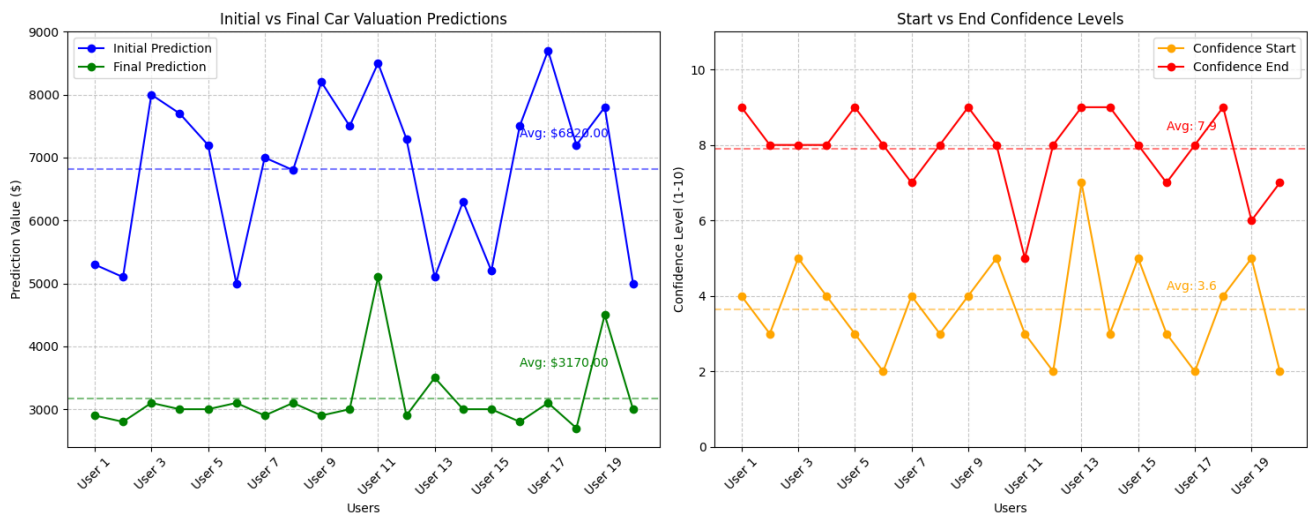
[Confidence Rating]

[Insert table of descriptive statistics]

Measure	Value
Mean Confidence Start	3.65
Mean Confidence End	7.90
Mean Confidence Change	+4.25

[Insert statistical test results] A paired-sample t-test revealed a statistically significant increase in participants' confidence ratings from pre- to post-explanation,  $p = 1.8311704 \times 10^{-10}$  ( $p < 0.001$ ).

[Optional: Insert visualization]



Key finding: Participants reported significantly higher confidence in their valuation estimates after receiving AI explanations, indicating that the prototype-based approach improved user self-assurance and trust in the decision process.

Enter your answer here

## Summary of Findings:

Hypothesis Supported?

H1	Yes	Participants' valuation predictions significantly shifted toward more accurate estimates after see
----	-----	--



H2

Yes

Participants' confidence levels significantly increased after receiving prototype-based explanation

**Discussion:** [75-100 words interpreting results in relation to complementary performance]

The results support that human-AI collaboration led to improved complementary performance. Many participants chose to align their final estimates closely with the AI model's prediction, often adopting the suggested value directly. Users reported that the AI explanations not only helped validate their instincts (e.g., confirming that high mileage reduces car value) but also enhanced their understanding of the magnitude of specific factors on pricing, which they previously found difficult to quantify. Additionally, participants who fully adopted the model's prediction (e.g., final estimate of 3000) tended to report slightly higher confidence ratings, suggesting that clear, prototype-based explanations strengthened both decision accuracy and user trust.