**User Experience and Artificial Intelligence Assignment1**

**2025712707 HongSeokjoo**

**Summary**

This paper explores how intelligent agents should apologize after making mistakes to effectively restore user trust. Specifically, it examines how an agent’s appearance (human-like vs. machine-like) and apology style (internal vs. external attribution) affect trust repair.

In a stock investment game experiment, participants were exposed to four conditions. Results showed that machine-like agents restored trust better with external attribution, while human-like agents were more trusted when they acknowledged their own fault.

These findings suggest that people do not always perceive AI as just a machine. Sometimes, they apply human-like expectations depending on how the agent looks and communicates. Therefore, the design of AI agents should match the apology strategy with the agent’s appearance. A sincere apology works better for human-like agents, while machine-like agents may benefit from explaining external circumstances.

**Critical Reflections**

While this study offers valuable insights into trust repair strategies for AI agents, there are some limitations worth noting.

First, the experiment was conducted in a relatively artificial setting using a short-term stock investment game. This limits the generalizability of the findings to real-world contexts, where human–AI relationships develop over longer periods and in more emotionally complex scenarios.

Second, the study focused only on the attribution type of apologies (internal vs. external), but did not consider other important aspects of how apologies are delivered—such as tone, emotional expression, or perceived sincerity. In real interactions, these factors can significantly influence how users respond to apologies.

Third, the manipulation of anthropomorphism relied on visual cues and self-introductions, assuming a uniform perception across participants. However, people may interpret the same agent very differently based on personal experiences or expectations. This individual variability in perceiving "human-likeness" was not fully accounted for.