Assignment 1, WASP Software Engineering Course Module 2023

Sygkounas Alkis, Örebro University

Project Title: Learning from Virtual and Real Worlds with Humans in the Loop

"Machine learning systems are a rapidly growing field aimed at enabling machines to interpret and understand data from their surroundings. One of the key challenges in this field is developing models that generalize well to new, unknown environments. This is crucial as real-world data can be limited, biased, or simply unavailable. To address this, a promising approach is to train models using both real and synthetic data into a hybrid learning, leading to improved generalization performance and reduced reliance on real-world data. Another promising development is the integration of human input in machine learning, known as "human in the loop" learning. By incorporating human feedback, machine learning algorithms can better understand and interpret visual data, leading to improved accuracy and performance. This approach has the potential to transform the way machines learn, bridging the gap between artificial and human intelligence. Overall, incorporating both real and synthetic data with human input is a promising strategy for improving machine learning models.

The current research in the autonomous driving domain uses machine learning to help vehicles drive autonomously. We currently have established a baseline of basic autonomous driving and we try to improve the training and decision making of the model by using the 'human in the loop' approach. In this phase, if the car makes an error, human experts step in to give real-time feedback. This synergy between human expertise and machine learning model is designed to improve the vehicle's decision-making process."

2.1. **State Transition Testing:**
State Transition Testing is based on the observation of system behavior for different input conditions. In simpler terms, it revolves around defining various states of the system and then testing the transitions between these states.
How it relates to my project: An autonomous driving system operates through numerous states. Consider states such as "idle", "accelerating", "cruising", "turning", or "emergency stop". Each of these states corresponds to specific behaviors of the car, and transitions between states are triggered by certain conditions or inputs. For instance, the state might transition from "cruising" to "emergency stop" when an unexpected obstacle is detected. In the context of the project, where we incorporate human feedback, the transitions become even more complex. A human might intervene to correct a perceived error, like when the car is about to make a wrong turn. The system must recognize this intervention and transition from its current state (e.g., "turning left") to the desired state (e.g., "continuing straight") seamlessly.
Why it's relevant: The dynamism of autonomous driving makes flawless state transitions essential. With the integration of human feedback, there's an added layer of complexity, making it even more crucial to ensure transitions are glitch-free. By meticulously testing these state transitions, especially after human intervention, we can pinpoint areas where the system might falter, ensuring the vehicle's decisions are both accurate and safe in real-time conditions.

2.2. Use Case Testing: Definition:
Use Case Testing focuses on determining the system's behavior for specific scenarios or 'use cases'. It involves identifying real-world situations the system might encounter and testing its responses.
How it relates to my project: Autonomous driving operates in an environment filled with countless scenarios. These can range from everyday occurrences like stopping at a red light to more challenging ones like

maneuvering around unexpected roadwork or reacting to a child running onto the street. The project's unique aspect is the human in the loop approach. This means that for each use case, not only should the car's initial response be tested, but also how the system adjusts when a human expert intervenes. For example, if the car decides to change lanes due to a perceived obstacle, but a human expert intervenes suggesting a different maneuver, the system's adaptability and responsiveness to this feedback should be put to the test.
Why it's relevant: Real-world application is the ultimate test for any autonomous system. The theoretical and synthetic scenarios can provide a foundation, but the car will be operating in a world that's unpredictable. By simulating these real-world scenarios and integrating human feedback, use case testing can provide invaluable insights. It ensures the machine not only understands these situations but also adapts and refines its decisions effectively based on human expertise.

3.1. Sentiment Analysis with AutoML

Sentiment Analysis leverages natural language processing (NLP) and computational linguistics to determine the sentiment or emotion conveyed in a text, typically categorizing it as positive, negative, or neutral. AutoML (Automated Machine Learning) is the process of automating the end-to-end process of applying machine learning to real-world problems, making it easier to apply and optimize machine learning models, including sentiment analysis.
How it relates to my project: While at first glance, sentiment analysis might seem unrelated to autonomous driving, its application can be found in analyzing feedback or interactions from the "human in the loop." As we incorporate real-time human feedback to correct and improve the car's decisions, sentiments from verbal or written feedback can be processed to understand the urgency, positivity, or negativity of the interventions. For instance, if we were collecting verbal feedback from testers or written reports post-test drives, sentiment analysis could help in quickly categorizing feedback into areas that need immediate attention (negative sentiments) versus areas that are working well (positive sentiments). AutoML would facilitate rapid deployment and optimization of this sentiment analysis, allowing for efficient feedback loops in the development process.
Why it's relevant: When we incorporate human feedback in our 'human in the loop' system, efficiently processing and understanding this feedback is paramount. For instance, suppose after a testing session, a human expert provides feedback such as, "The car's response in rainy conditions was unnerving and felt dangerous." A sentiment analysis can quickly categorize this as a negative sentiment and highlight it as a priority area for improvement. With AutoML, we can streamline this process, deploying optimized sentiment analysis models rapidly. This ensures we prioritize adjustments based on the emotional urgency conveyed by the experts, allowing our autonomous system to adapt in real-time and ensuring our vehicle's performance meets human expertise and expectations.

3.2. Boundary Value Testing for Traditional and ML Software with DevBots AI/ML support: Boundary Value Testing is a software testing technique that involves determining the edge or boundary conditions and testing them. It's often used to catch edge case errors. DevBots with AI/ML support are automated systems/tools that assist developers in various tasks, including testing.
How it relates to my project: Given the vast complexity of autonomous driving and the numerous variables at play, edge cases are abundant. For instance, while a car might perform well in standard driving conditions, how it reacts at the boundary conditions, like during extreme weather or when faced with unexpected road obstructions, is crucial. The human in the loop approach we've integrated can provide feedback during these boundary conditions, and DevBots with AI/ML support can assist in automating the process of identifying, testing, and refining the system's response to these conditions. For example, if during testing, a human expert intervenes when the car misjudges a boundary condition (like stopping distance on a wet road), the DevBot could automatically log this, retest under similar conditions, and validate improvements based on AI/ML algorithms.

Why it's relevant: It's imperative for our autonomous vehicles to be adept not only in common scenarios but also in rare and extreme ones. For example, while our car might be calibrated to handle standard lane changes, how it behaves when confronted with a sudden lane merge due to construction or an unexpected obstruction is essential. Boundary Value Testing allows us to focus on these edge scenarios. If, during testing, a human expert flags an issue during such an edge scenario, our DevBots, supported by AI/ML, could log this, automatically simulate similar conditions, and verify any system refinements. Through this, we ensure comprehensive refinement of our vehicle's decision-making processes, safeguarding robust performance across a myriad of driving conditions.

4.1. Human-Computer Interaction (HCI)

Human-Computer Interaction (HCI) deals with the design and use of computer technology, centered on the interfaces between humans (users) and computers. This field seeks to understand the interactions between users and computers and to design technologies that let those interactions be more effective.

Research Challenges:

1. Adaptive Learning Systems: Challenge: Creating AI models that not only incorporate human feedback but can adapt in real-time without requiring periodic retraining sessions. How can an AI system learn continually and effectively from human feedback without negatively affecting its prior knowledge?
2. Personalized Driving Experience: Challenge: Determining the extent to which autonomous vehicles should adapt to individual driving behaviors. Should there be a limit to personalization, considering safety standards and shared road responsibility?
3. Real-time Intervention Analytics: Challenge: Effectively analyzing human interventions in real-time and integrating the feedback immediately. This would require a deep understanding of real-time data processing and instantaneous model adaptation.
4. Validating Feedback Authenticity: Challenge: Ensuring that the feedback provided by the human driver is authentic and safe. How can we validate that interventions made by the human driver are the best responses in those scenarios?
5. Interface Dynamics: Challenge: Building an intuitive, anticipative, and interactive interface for drivers that aligns with their expectations and reduces cognitive load, especially during critical interventions.
6. Data Privacy and Security: Challenge: With rich datasets detailing various intervention scenarios, maintaining the privacy and security of this data becomes paramount, especially if used commercially.

Commercial Opportunities:

1. Customizable Autonomous Vehicles:
Personalized Driving Experience: With the ability to adapt based on individual human feedback, autonomous vehicles can offer a driving experience tailored to each user's preferences. This can lead to a range of customizable car models, where users can choose a vehicle based on its adaptability level, creating a niche market.

2. Advanced Driving Assistance Tools:
Real-time Advisory Systems: By harnessing feedback and its subsequent analysis, commercial opportunities arise for tools that offer real-time advice to the driver. This system would highlight potential areas of concern, possible interventions, or alternative route suggestions, acting as an advanced driver-assistance system (ADAS).

3. Training Platforms for Transitioning Drivers:
Simulated Training Modules: Given the feedback from human interventions, there's a commercial possibility for creating simulation-based training modules. These would help drivers transitioning from traditional vehicles to autonomous ones, offering them scenarios where they're likely to intervene and guiding them on best practices.

4. Enhanced In-vehicle Interfaces:
Upgrade-able Interface Modules: Considering the context-aware controls and dynamic learning, there's a commercial prospect for developing upgradeable interface modules. These would enhance the user interface over time, based on frequent intervention scenarios and user feedback, offering drivers a continually improving interaction experience.

Application in AI/ML:

1. **Enhanced Feedback Mechanisms**:
Direct Annotation on Feedback: Post any driving intervention, the human operator can directly annotate the event, highlighting the reason for the intervention. An intuitive interface would allow them to choose from predefined categories like 'Traffic Anomaly', 'Safety Concern', 'Environmental Conditions', or even manually enter specifics.
T**emporal Clustering of Interventions**:
Machine learning algorithms can cluster similar intervention scenarios based on timing, location, and the nature of the interventions. This clustering helps identify patterns where human intervention is frequent, signaling areas where the AI system requires refinement.
2. Adaptive Human-Computer Interfaces (AHCI):
Context-Aware Controls: Depending on the scenario, the interface of the vehicle can adapt to show the most relevant controls or information. For instance, in dense urban settings, the interface could provide a quick overview of surrounding vehicles, pedestrians, and traffic signals, given these are common reasons for human interventions in such scenarios.
Predictive System Alerts: By analyzing past interventions and their reasons, the system can preemptively alert the human operator of potential upcoming situations where they might need to intervene. This is not about predicting exact interventions but more about heightening driver awareness in familiar intervention scenarios.
3. Dynamic Learning from Interventions:
Immediate Model Feedback: Post-intervention, the AI system not only collects feedback but also contrasts its own decision-making with the human operator's actions. This contrast serves as a powerful training data point, emphasizing areas where the model's predictions deviated from human intuition.
4. Validation and Testing based on Human Interventions:
Scenario Replication in Simulated Environments: Frequently occurring human intervention scenarios can be replicated in a simulated environment. This allows the AI model to be tested against these scenarios repeatedly, ensuring it learns to handle them more proficiently over time.

4.2: Quality assurance:
Quality Assurance (QA) ensures software aligns with standards and requirements, identifying and correcting defects. For AI-driven systems like autonomous vehicles, QA's role is heightened due to the integration of "human in the loop" feedback, demanding rigorous validation of both AI models and human interactions.

Research Challenges:

1. Real-time Model Verification: Challenge: Establishing a system that can validate real-time changes made to the driving model based on human feedback. Ensuring that every tweak to the AI due to intervention is reliable, and doesn’t introduce new errors.
2. Consistent Feedback Integration: Challenge: Human drivers may provide varying feedback for similar driving scenarios. QA must ensure that this feedback does not lead to inconsistent or unpredictable model behavior.
3. Comprehensive Test Environments: Challenge: Designing test environments that emulate real-world driving scenarios is crucial. This is especially challenging when trying to anticipate and recreate unique or rare driving situations where human intervention might be critical.

4. Feedback Prioritization: Challenge: With the continuous influx of human feedback, determining which feedback should be immediately integrated versus what should be queued or further evaluated becomes critical for maintaining model quality.

Commercial Opportunities:

1. QA Consultancy Services: Given the intricacies of integrating human feedback, there's a potential market for firms specializing in offering QA services tailored to "human in the loop" models, guiding companies on best practices and methodologies.
2. Feedback Integration Tools: Commercial software tools that can seamlessly integrate and prioritize human feedback into machine learning models while maintaining version control, ensuring consistent and quality updates.

Application in AI/ML:

Predictive Analysis for Feedback Integration: Machine learning can be employed to analyze patterns in human feedback. By recognizing consistent interventions by humans in specific scenarios, the system can be trained to anticipate and even correct its actions before human intervention is needed. This iterative feedback loop will progressively reduce the frequency of human interventions.
Anomaly Detection in Driving Patterns: Using unsupervised learning techniques, the system can monitor driving patterns and detect anomalies. If a particular action deviates significantly from established norms (based on both AI decisions and human interventions), it can be flagged for review. Simulated Environment Testing: Deep reinforcement learning can be employed to train autonomous vehicles in simulated environments. These environments can recreate real-world scenarios where human interventions were frequent, allowing the AI to learn in a risk-free setting.
Real-time Model Adjustment: Incorporating online learning strategies, where the AI model can adjust its parameters in real-time based on the feedback without requiring extensive retraining. This ensures that the vehicle's AI system remains adaptive and up-to-date, especially in dynamic driving conditions.