



TECHNOLOGY FEATURE Industrial News / Technology Report / Research Communication

AI to the Rescue: Preventing Toxicity in Online Games

Online games are a shared experience for millions, but toxicity is on the rise Will AI bring justice and respect to the game?.

By Muhammad Muneeb | Student / Tech Enthusiast | Published: June 2026

m7349197@gmail.com | Position: Aspiring computer scientist

Introduction

Online games have become more than entertainment, a digital phenomenon. Whether playing games or completing quests, gamers engage in real-time with others from around the world. Games such as PUBG and Call of Duty have amassed huge communities, including a burgeoning community in Pakistan.

But with this expansion comes a problem - toxicity. Swastikas, bullying, swearing and cheating are all too prevalent in the gaming world. This can make the experience for some, particularly beginners, frustrating. So, can technology - specifically artificial intelligence - help to address an issue that has been generated in a digital environment?

In Pakistan's burgeoning gaming and esports scene, it is about more than enhancing the gaming experience. It is about fostering safe and inclusive spaces that promote engagement, learning and careers.

"Playing games is supposed to be a test of your ability and not your sanity, but toxicity can make that difficult."

Technology Overview

AI is increasingly being applied to moderating online gaming. AI is able to monitor gameplay in real time, rather than relying solely on reports from players. This technology can pick up on offensive language and harassment, as well as even harmful voice messages.

For instance, voice recognition technologies can monitor gameplay conversations and detect inappropriate language or hostile language. Likewise, AI text filters can scan chat data in real time and prevent or alert of potential abuse. Other techniques involve awarding players with behaviour points for positive interactions and deducting points for negative behaviours.

The other key innovation is predictive moderation. Rather than respond to problems as they arise, AI models can predict which players are at high risk for engaging in toxic behavior. This enables game companies to proactively intervene, by restricting communication options or sending warnings.

Such innovations will be beneficial in Pakistan, where online gaming is gaining momentum. Children often lack online etiquette when they start playing online games. AI-driven solutions can help to educate and keep a level playing field.



Key Findings & Impact

Surveys indicate that negative player behaviour is a major cause of player drop-off. Toxicity decreases retention rates and negatively impacts gaming platforms' reputation.

AI moderation has already been successful. Games using intelligent content filtering have less harassment and higher player satisfaction. Instant detection means that offensive content is dealt with quickly, minimising its effects.

But AI is far from perfect. Context understanding is a key challenge. What may be offensive in one context may not be in another and even across cultures. Given the language barriers for Pakistani gamers (Urdu, English, and many local languages), this is even more challenging.

Then there's the issue of fairness. Excessive dependence on automation can result in unfair penalties, where players are penalised unfairly. So there is a need for a hybrid approach that leverages the efficiency of AI while allowing for human intervention.

Despite the risks, there are many advantages. Improving the gaming culture can enhance Pakistan's esports scene, attract new gamers and open up gaming careers. And it can foster a healthy online culture among young people, which can carry over to all online activities.

What's Next?

The key to more effective gaming moderation is more intelligent AI. With advancements in machine learning, AI will have a greater grasp of context, tone and cultural subtleties. This will enable more effective and less obtrusive moderation.

Developers need to prioritise community, as well as tech. Incentivising good behaviour, raising awareness and educating gamers can be added to AI approaches.

Pakistan can take this chance to establish a positive gaming culture. Schools, gamers and game developers can collaborate to encourage positive online behaviour and responsible gaming.

Ultimately, AI can only help with toxicity - it can't stop it. It's up to the players to make a difference. It takes more than code to create a safe space for gaming, it takes action by the players themselves.

About the Author

Muhammad Muneeb is an enthusiastic student and future software engineer, who has a keen interest in artificial intelligence and communities in the digital world. He is interested in the impact of technology on human behaviour, and writes to educate people about new tech issues. You can contact him on m7349197@gmail.com.

Keywords: Artificial Intelligence; Online Games; Cybersecurity; Online Communities; Pakistan Tech

References

- [1] Blackburn, J., & Kwak, H. (2014) Shut up noob!: Predicting Crowdsourced Decisions on Toxic Behavior in Online Games.
- [2] Latest Developments in AI Moderation, International Journal of Computer Science, (2015).