Google Summer of Code 2025 Proposal

Proposal Title:

Automated Rule Suggestion and Enhanced Reporting for Checkstyle

Abstract:

This project aims to expand the capabilities of Checkstyle by introducing two major features:

- Automated Rule Suggestion: A module that uses basic machine learning techniques to analyze a codebase and recommend new Checkstyle rules based on common stylistic patterns or inconsistencies found across the project. This feature can accelerate Checkstyle adoption for new teams and reinforce best practices.

- Enhanced Reporting System: A more informative and developer-friendly reporting system thatuses contextual code snippets, improved categorization of warnings, and visual aids to make rule violations more understandable. This can help teams reduce technical debt faster and improve their code quality iteratively.

Benefits to the Community:

Checkstyle is widely used for enforcing coding standards in Java projects. However, creating custom rules or understanding rule violations can be daunting, especially for newcomers. By adding:

- Rule Suggestion: Developers can bootstrap their configuration using intelligent suggestions,reducing manual effort.

- Enhanced Reports: Teams get faster feedback and better insights, making Checkstyle moreusable in CI pipelines and educational tools.

These improvements will make Checkstyle more accessible and significantly increase its usability, adoption, and contribution potential.

Deliverables:
Community Bonding Period:
- Familiarize with Checkstyle's architecture and rule engine
- Engage with mentors and contributors
- Finalize exact implementation scope and evaluation criteria

Phase 1: (Weeks 1-6)

- Design and prototype the rule suggestion module
- Train a lightweight ML model using frequency-based heuristics or pattern mining (no external data)
- Implement an interface to review and accept rule suggestions- Begin designing the enhanced reporting format (HTML/JSON) - Midterm evaluation deliverables:
- Working prototype of rule suggestion engine
- Initial mockup of the enhanced report system

Phase 2: (Weeks 7-12)

- Finalize and optimize rule suggestion module
- Implement enhanced report generation with visual improvements
- Add context-aware reporting using code snippets
- Write documentation and usage guides- Final evaluation deliverables:
- Complete integration into Checkstyle core or plugin
- Test coverage and configuration samples
- User guide for both features

Timeline:

- May 20 - June 16 (Community Bonding): Study Checkstyle codebase, finalize proposal withmentors
- June 17 - July 29 (Phase 1): Build rule suggestion engine, basic report redesign
- July 29 - Aug 26 (Phase 2): Finalize enhanced reporting, polish all features, testing
- August 26 - September 2: Submitting final work, writeup, code cleanup

---

Technologies:

- Java (core language)
- ANTLR (for parsing)
- Gradle (build tool)
- Git (version control)
- JSON/XML (report formatting)
- Basic Machine Learning concepts (pattern analysis, frequency heuristics)

Why Me?

I am a passionate developer with solid experience in Java and static code analysis. I've contributed to open-source projects and built custom tooling in past internships. I enjoy blending software engineering with intelligent systems, and this project matches both my interest and skillset.

Relevant Skills:

- Java: 3+ years of experience
- Familiarity with parsing tools and rule engines
- Git & CI: Used extensively in team projects
- Strong understanding of developer pain points regarding code quality

Long-Term Vision:

I hope to continue contributing to Checkstyle even after GSoC. I aim to maintain and improve the modules I create and possibly assist others onboarding with documentation and mentoring.

Personal Information:

Name: T. Rohit Chary

Email: Rohittukkapuram@gmail.com

University: MVSR Engineering College

Degree Program: B.Tech in Electronics and Communication Engineering

Year of Study: Fourth Year

Country: India