**Functional & Performance Testing Template**

**Model Performance Test**

|  |  |
| --- | --- |
| **Date** | **27 June 2025** |
| **Team ID** | **LTVIP2025TMID22707** |
| **Project Name** | **SmartSDLC: AI-Enhanced Software Development Lifecycle** |
| **Maximum Marks** | **2** |

**Test Scenarios & Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **🧪 Test Scenario** | **Input** | **Expected Behaviour** | **Actual Result** | **Status** |
| **1. Grammar Correction** | Paragraph with multiple grammatical errors | Output corrected paragraph + human-readable explanations per correction | Corrected content, explanations match severity of errors | ✅ Pass |
| **2. Spelling Fixes** | Text with spelling mistakes (e.g., “recieve”, “definately”) | Correct spelling in output; explanations note original typos | Spelling corrected; explanations precise | ✅ Pass |
| **3. Multilingual Detection** | Short text each in English, Spanish, French, Hindi, Chinese, Arabic | Output shows detected language + confidence score via langid | Detected languages with ≥90% confidence | ✅ Pass |
| **4. Multilingual Grammar Practice** | User picks Spanish; incorrect subject-verb agreement | Exercise prompt, answer input, checks correctness, provides examples + answer key | Grammar quiz delivered, feedback accurate | ✅ Pass |
| **5. Sentence Formation Exercise** | Jumbled words in French: “chien le mange” | Correct ordered sentence; explanation of order | Output “le chien mange”; clear syntax explanation | ✅ Pass |
| **6. Verb Tense Drill** | German present-tense prompt; user answers in past tense | Model's assessment, explanation, corrected answer & key | Identified tense error reliably | ✅ Pass |
| **7. Text Analysis Metrics** | Long English text (~200 words) | Outputs word count, avg word length, sentence complexity | Metrics match independent analyzer | ✅ Pass |
| **8. Visual Bar Chart** | Analysis result from Test 7 | Bar chart visualizing metrics | Chart rendered correctly via Gradio + Matplotlib | ✅ Pass |
| **9. Radar Chart Competency** | User writing sample in Hindi | Radar chart output across dimensions (grammar, vocabulary, structure) | Radar visual displays with correct scaling | ✅ Pass |
| **10. Language Detection Error** | Mixed-language paragraph | Langid returns top language; warns if low confidence (<70%) | Shown correct top language; confidence displayed | ✅ Pass |
| **11. Real-time Correction Latency** | Continuous user typing | < 2 sec response time per query | ⚠️ Averaged ~1.8s on Colab with cached model; slight delay on longer input |  |
| **12. Model Loading & Cold Start** | Fresh session without cache | First prompt ~10s; subsequent <2s | ✅ Pass |  |
| **13. Handling Edge Cases** | Empty input, emojis only, gibberish strings | Graceful prompt: “Please enter valid text.” No crash | ✅ Pass |  |
| **14. Non-Latin Scripts** | Input in Arabic or Chinese | Detection + correct spelling/grammar handling | Works accurately for Chinese; Arabic minor misclassification (85% vs 92%) | ⚠️ Partial |
| **15. Explanations Quality** | Complex corrections | Explanations are concise, correct, informative | Observed occasional overly technical phrasing | ⚠️ Partial |