***Classification of Events***

**True Positives (TP) - Correct Predictions by LLM**

These are events that the LLM correctly predicted and also happened during the real Battle of Gettysburg.

| **Event** | **Definition** |
| --- | --- |
| The battle occurred at a strategically important location | Gettysburg was a major transportation hub, making it valuable for both Union and Confederate forces. |
| The defending force (Union) held strong defensive positions | The Union Army occupied high ground positions (Cemetery Hill, Little Round Top), providing a significant advantage. |
| The attacking force (Confederates) had early success but was ultimately repelled | The Confederates achieved early gains but failed in key assaults, especially Pickett’s Charge. |
| Artillery played a crucial role in the battle | Both sides used extensive artillery bombardments, particularly before Pickett’s Charge. |
| Cavalry engagements influenced the battle | Union and Confederate cavalry skirmishes shaped reconnaissance and flank security. |
| The battle lasted multiple days with momentum shifts | The battle lasted three days (July 1–3, 1863), with fluctuating control over key positions. |
| Logistics and supply chains impacted the battle | The Union had stronger supply lines, while the Confederates faced resupply issues due to their extended campaign. |
| The battle significantly impacted the broader war | Gettysburg was a turning point in the Civil War, halting Confederate momentum. |
| The defending side (Union) won the battle | The Confederate Army was forced to retreat after failing to break Union lines. |
| The battle resulted in massive casualties | Over 50,000 soldiers were killed, wounded, or missing, making it one of the bloodiest battles in U.S. history. |

Total True Positives (TP): **10**

**False Positives (FP) - Incorrect Predictions by LLM**

These are events that the LLM predicted but did NOT actually happen in the real Battle of Gettysburg.

| **Event** | **Why It Did NOT Happen?** |
| --- | --- |
| The defending side launched a large-scale counteroffensive after the battle | The Union won but did NOT aggressively counterattack Lee’s retreating forces. |
| Foreign nations (Britain, France) directly intervened | Britain and France never officially entered the war, despite Confederate hopes for recognition. |
| The defending army suffered severe supply shortages | The Union was well-supplied throughout the battle. |
| The attacking force (Confederates) relied heavily on naval power | This was a land battle; no naval forces were involved. |
| Urban warfare played a major role | Gettysburg was not a major city; most fighting occurred in open fields, ridges, and farmland. |
| The battle ended in a ceasefire rather than a retreat | The Confederates retreated after their defeat; no formal ceasefire was declared. |
| Guerrilla warfare tactics significantly shaped the battle | Both sides fought in conventional battle lines, not irregular guerrilla tactics. |
| The attacking force primarily consisted of foreign mercenaries | The Confederate Army was composed of Southern volunteers, not hired mercenaries. |
| A major river crossing played a decisive role in the battle | The battle took place far from major rivers, though Lee’s retreat was later affected by a swollen river. |
| Chemical weapons or other unconventional warfare were used | The battle was fought with conventional 19th-century weapons, with no chemical warfare. |

Total False Positives (FP): **10**

**False Negatives (FN) - Missed Predictions by LLM**

These are real events that happened during the Battle of Gettysburg but were NOT predicted by the LLM.

| **Event** | **Definition** |
| --- | --- |
| The battle began unexpectedly as Confederate forces searched for supplies | The battle was not pre-planned; it started when Confederate troops encountered Union cavalry while searching for shoes. |
| Pickett’s Charge was a decisive moment | The Confederates' final assault (Pickett’s Charge) was a key turning point, leading to their defeat. |
| Union forces used interior lines for rapid reinforcement | The Union’s compact positioning allowed for faster movement and reinforcement of key areas. |
| General Meade was the Union commander | The LLM did not identify General George G. Meade, who commanded Union forces. |
| General Lee took responsibility for the Confederate defeat | After the battle, Robert E. Lee admitted his failure and offered to resign (which was declined). |
| The Confederate retreat was slowed by a supply train and wounded soldiers | The Confederates struggled to retreat efficiently due to large numbers of wounded and supply wagons. |
| Heavy rain complicated the Confederate retreat | A storm slowed Confederate forces as they withdrew from Pennsylvania. |
| Gettysburg Address followed the battle | Abraham Lincoln’s Gettysburg Address, given later that year, became one of the most famous speeches in U.S. history. |

Total False Negatives (FN): **8**

***Final Results***

| **Metric** | **Value** |
| --- | --- |
| Total True Positives (TP) | **10** |
| Total False Positives (FP) | **10** |
| Total False Negatives (FN) | **8** |
| Precision | **50.00%** |
| Recall | **55.56%** |
| F1-Score | **52.63%** |

***Critical Analysis of LLM’s Prediction***

1. Major Issue: The LLM Predicted the Wrong Winner

* The LLM incorrectly assumed the attacking force (Confederates) won, when in reality, the Union (defending force) won the battle.
* The Confederates launched Pickett’s Charge, failed, and retreated, ending their invasion of the North.

2. Strengths in LLM’s Predictions

Correctly identified major aspects of the battle:

* Gettysburg was a turning point in the war.
* The Union held strong defensive ground.
* Artillery and cavalry played key roles.
* High casualties on both sides.
* The battle had lasting political consequences.

3. Weaknesses in LLM’s Predictions

Overestimated aggressive Union counterattacks:

* The Union did not pursue Lee aggressively after his retreat.
* The Confederacy was NOT entirely crushed, just forced to retreat.

Missed key tactical moments:

* Pickett’s Charge was central to the battle outcome.
* The battle started unexpectedly over supplies.

Invented historical inaccuracies:

* Naval power played no role.
* Guerrilla warfare was not a factor.
* No foreign intervention.

***Final Verdict***

* The LLM’s prediction was weak (~52.63% F1-score) due to incorrect outcome prediction and multiple historical inaccuracies.
* Improving recall (identifying more real-world details) and correcting historical errors would significantly improve future assessments.