***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 took place in a strategically important town | Gettysburg was a critical transportation hub, making it valuable for both Union and Confederate forces. |
| The defending force held strong defensive positions | The Union Army occupied high ground (e.g., Cemetery Hill, Little Round Top), providing a defensive advantage. |
| The attacking force was initially successful but later repelled | The Confederates achieved early gains but failed in key assaults, especially Pickett’s Charge on the final day. |
| 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 shifts in momentum | The battle occurred over three days (July 1–3, 1863), with fluctuating control of different areas. |
| Logistics and supply chains influenced the battle outcome | The Union had stronger supply lines, while the Confederates struggled with resupply. |
| The battle had a major impact on the broader war | Gettysburg marked a turning point in the Civil War, weakening Confederate momentum. |
| The defending side won the battle | The Union defeated the Confederate invasion, forcing a retreat to Virginia. |
| High casualties were suffered by both sides | The battle resulted in over 50,000 combined casualties, making it one of the bloodiest conflicts in U.S. history. |
| The battle led to a famous speech that shaped national identity | Abraham Lincoln’s Gettysburg Address emphasized unity and democracy, reinforcing the Union’s resolve. |

Total True Positives (TP): **11**

**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 intervened in the battle | Neither Britain nor France intervened despite Confederate hopes for recognition. |
| The defending army suffered from critical supply shortages | The Union Army was well-supplied, unlike the Confederates, who faced shortages. |
| The attacking force relied heavily on naval power | This was a land battle with no naval involvement. |
| Urban warfare played a major role | Fighting mostly occurred in open fields, ridges, and farmland, with minimal urban combat in Gettysburg itself. |
| The battle ended in a ceasefire rather than a retreat | The Confederates retreated after their defeat; there was no negotiated ceasefire. |
| Guerrilla warfare tactics significantly shaped the battle | Both sides fought in conventional battle lines, with no significant guerrilla tactics. |
| The defending side used large numbers of war elephants | This is a clear historical inaccuracy—war elephants were not used in the American Civil War. |
| The attacking force was primarily composed of mercenaries | The Confederate Army was largely composed of volunteers from the Southern states, not hired mercenaries. |
| A major river crossing played a decisive role in the battle | The battle took place far from major river crossings, though retreating Confederates later struggled with river conditions. |

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 by accident 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 defining moment of the battle | The Confederates’ final, disastrous assault (Pickett’s Charge) was a key turning point in the Union victory. |
| Union forces used interior lines for rapid reinforcements | The Union Army used its compact positioning to move reinforcements faster than the Confederates. |
| General Meade was the Union commander | The LLM did not identify General George G. Meade, who commanded Union forces at Gettysburg. |
| General Lee took full responsibility for the Confederate defeat | After the battle, Robert E. Lee admitted his failure and offered his resignation (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 after they began their withdrawal from Pennsylvania. |

Total False Negatives (FN): **7**

***Final Results***

| **Metric** | **Value** |
| --- | --- |
| Total True Positives (TP) | **11** |
| Total False Positives (FP) | **10** |
| Total False Negatives (FN) | **7** |
| Precision | **52.38%** |
| Recall | **61.11%** |
| F1-Score | **56.48%** |

**Critical Analysis of LLM’s Prediction**

1. Major Issue: The LLM Predicted the Wrong Winner

* The LLM incorrectly assumed the attacking force won, when in reality, the Union (defending force) won the battle.
* The Confederate Army was forced to retreat, ending its 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, cavalry, and logistics played crucial 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 war elephants or mercenaries were involved.

***Final Verdict***

* The LLM’s battle prediction was weak (~56.48% F1-score) due to overestimating Union aggression, missing key events, and predicting the wrong winner.
* Improving recall (identifying more real-world details) and correcting historical inaccuracies would significantly improve future assessments.

Τέλος φόρμας