***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 Midway.

| **Event** | **Definition** |
| --- | --- |
| The battle was a major naval engagement between aircraft carriers | Midway was primarily a carrier-based battle, with aircraft playing the decisive role in the outcome. |
| The defending side (USA) had an intelligence advantage | The U.S. broke Japanese codes, allowing them to anticipate and prepare for the attack. |
| The attacking force (Japan) aimed to seize a strategic island | Japan's goal was to capture Midway Atoll, forcing the U.S. into a weakened position in the Pacific. |
| The defending side (USA) set up an ambush | The U.S. used deceptive tactics and pre-positioned its carriers to counter Japan’s attack. |
| Aircraft played the primary offensive role | The battle was won through carrier-launched dive bombers and torpedo planes, not direct ship-to-ship combat. |
| The attacking force (Japan) had superior numbers at the start | The Japanese Navy had more aircraft carriers and battleships at the outset of the battle. |
| The defending side (USA) had a home-field advantage | U.S. forces operated from Midway Atoll and nearby airbases, giving them strategic positioning. |
| The defending side (USA) suffered early losses but recovered | The battle began with U.S. forces struggling, but they later turned the tide with well-timed strikes. |
| Multiple Japanese aircraft carriers were destroyed | The U.S. sunk four Japanese carriers (*Akagi, Kaga, Soryu, and Hiryu*), crippling Japan’s naval power. |
| The battle shifted the balance of power in the Pacific | Midway marked a turning point in World War II, putting Japan on the defensive. |

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

F**alse Positives (FP) - Incorrect Predictions by LLM**

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

| **Event** | **Why It Did NOT Happen?** |
| --- | --- |
| The defending side (USA) launched a major counter-invasion after victory | The U.S. did not immediately invade Japanese-held territory; they remained defensive for months after Midway. |
| Japan successfully landed troops on Midway Atoll | The Japanese invasion force never landed, as their carriers were destroyed before they could launch an amphibious assault. |
| The battle involved a significant submarine attack | While submarines played a minor role, Midway was mostly an aerial battle. |
| A ceasefire or temporary truce was declared mid-battle | No diplomatic negotiations occurred; the battle was fought until Japan's fleet was forced to retreat. |
| A last-minute reinforcement changed the outcome | The battle was decided by existing forces, with no major reinforcements affecting the result. |
| Japan relied on radar technology for its fleet defense | Japan had limited radar capabilities, and its forces primarily relied on reconnaissance planes for situational awareness. |
| U.S. forces deployed kamikaze-style tactics | Kamikaze attacks were not used by the U.S.; Japan would later employ them in 1944–45. |
| The U.S. fleet suffered equal carrier losses as Japan | The U.S. only lost one carrier (USS Yorktown), while Japan lost four. |

Total False Positives (FP): **8**

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

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

| **Event** | **Definition** |
| --- | --- |
| The U.S. deliberately set a trap by broadcasting false information | The U.S. sent deceptive messages about Midway's water supply, tricking Japan into confirming its attack plans. |
| Japan’s initial scout planes failed to detect U.S. carriers in time | A crucial Japanese scout plane was delayed, preventing them from spotting U.S. forces early enough. |
| U.S. torpedo bombers suffered devastating losses before dive bombers struck | Early torpedo bomber squadrons were wiped out, but their distraction helped set up the decisive dive bomber attack. |
| Weather conditions affected visibility and engagements | The battle was influenced by cloud cover and visibility issues, impacting aircraft operations. |
| Admiral Yamamoto personally commanded the Japanese fleet but was distant from the battle | Yamamoto planned the battle but was too far away to respond effectively once the tide turned. |
| The Japanese fleet was divided, making coordination difficult | Japan’s main force and supporting ships were spread out, complicating command and delaying reinforcements. |
| U.S. aircraft had technological advantages in survivability and tactics | While some Japanese planes were superior, U.S. aircraft benefited from better defensive formations and coordination. |
| Midway was Japan’s last major offensive in the Pacific | After Midway, Japan never regained the strategic initiative and shifted to a defensive war. |
| Surviving Japanese ships retreated under cover of darkness | After their carriers were lost, Japanese forces used nightfall to escape further destruction. |

Total False Negatives (FN): **9**

***Final Results***

| **Metric** | **Value** |
| --- | --- |
| Total True Positives (TP) | **10** |
| Total False Positives (FP) | **8** |
| Total False Negatives (FN) | **9** |
| Precision | **55.56%** |
| Recall | **52.63%** |
| F1-Score | **54.02%** |

**Critical Analysis of LLM’s Prediction**

1. Major Issue: The LLM Predicted the Wrong Winner

* The LLM incorrectly assumed the attacking force (Japan) won, when in reality, the defending force (USA) decisively won the battle.
* Japan lost four aircraft carriers, shifting the balance of power in the Pacific.

2. Strengths in LLM’s Predictions

Correctly identified major aspects of the battle:

* Aircraft carriers were central to the battle.
* U.S. intelligence and codebreaking played a crucial role.
* Japan started with an advantage but was ultimately outmaneuvered.
* The U.S. suffered early setbacks but staged a comeback.
* The battle was a turning point in World War II.

3. Weaknesses in LLM’s Predictions

Overestimated Japan’s ability to sustain the battle:

* Japan’s logistical and strategic weaknesses were not properly accounted for.
* The U.S. suffered some early losses, but not enough to shift the outcome.

Missed key tactical moments:

* Japan’s scouting failures played a critical role in their defeat.
* The U.S. used deception tactics to lure Japan into the ambush.

Invented historical inaccuracies:

* No land invasion of Midway occurred.
* No ceasefire or truce was declared.
* No large-scale submarine attacks changed the outcome.

***Final Verdict***

* The LLM’s prediction was weak (~54.02% F1-score) due to incorrect outcome prediction and historical inaccuracies.
* Improving recall (identifying more real-world details) and correcting historical errors would significantly improve future assessments.