***Classification of Events***

Below is a **detailed definition** of each event and situation, categorized into **True Positives (TP), False Positives (FP), and False Negatives (FN)**.

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

These are events that both the LLM and the real Battle of Normandy had in common.

| **Event** | **Definition** |
| --- | --- |
| **Strong defensive fortifications** | The German **Atlantic Wall** was heavily fortified with bunkers, artillery, machine-gun nests, and obstacles on beaches to slow down the invaders. |
| **Elite troops and experienced leadership for defenders** | Germany deployed **Waffen-SS and Wehrmacht elite divisions** under **Field Marshals Rommel and Rundstedt** to defend against invasion. |
| **Fuel shortages weakened defenders** | Germany suffered **critical fuel shortages**, limiting tank and vehicle movement, hindering reinforcements. |
| **Defenders had disrupted supply lines** | The **Allied bombing campaign** (Operation *Transportation Plan*) targeted **railways, roads, and supply depots**, disrupting German logistics. |
| **Defenders misinterpreted the invasion site** | Hitler and the German High Command believed **the real attack would come at Pas-de-Calais**, delaying reinforcements for Normandy. |
| **Attackers had superior naval and air power** | The **Allies dominated the skies and seas**, preventing German air support and reducing their ability to counter landings. |
| **Attackers had strong logistical support** | The Allies set up **floating harbors (Mulberry Harbors)** and **underwater fuel pipelines (PLUTO)** to sustain operations. |
| **Multinational coordination of attackers** | The invasion involved **troops from the U.S., U.K., Canada, Free French, Poland, and others**, demonstrating coordinated efforts. |
| **Attackers faced difficult landing terrain** | Beaches were **heavily mined, had barbed wire, Czech hedgehogs, and wooden stakes** making landings extremely difficult. |
| **Weather conditions influenced the attack** | **D-Day was delayed from June 5 to June 6, 1944** due to bad weather; meteorological forecasting played a key role in the decision. |
| **Attackers suffered high casualties in some areas** | **Omaha Beach saw extreme losses** (over 2,000 American casualties) due to strong German resistance. |
| **Deception played a major role in misguiding defenders** | **Operation Fortitude** (a fake army under General Patton) tricked Germans into believing the invasion was at **Pas-de-Calais**. |
| **Defenders' delayed response led to failure** | Due to Hitler’s **hesitation and micromanagement**, German reinforcements were not immediately deployed. |
| **Air and naval dominance dictated the battle outcome** | Allied **bombing, air support, and naval shelling** softened defenses and prevented German counterattacks. |
| **Defenders struggled with morale** | Many **German conscripts (especially Ost Battalions)** were poorly trained and surrendered quickly. |
| **Supply chain collapse affected defenders** | The Germans **couldn’t reinforce fast enough** due to destroyed roads, bridges, and fuel shortages. |
| **Industrial and infrastructure damage after battle** | Allied bombings and battles left **Normandy in ruins**, crippling German logistics and war efforts. |
| **International pressure influenced decisions** | The invasion had **global political stakes**, and pressure was on the Allies to open a Western Front against Germany. |
| **War shifted in favor of the attacking coalition** | D-Day marked the beginning of **Germany’s decline**, leading to the eventual liberation of France and collapse of the Nazi regime. |

Total True Positives: **19**

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

These are **events predicted by the LLM that did NOT actually happen** in the Battle of Normandy.

| **Event** | **Why It Did NOT Happen?** |
| --- | --- |
| **Defenders launched large-scale counterattacks** | The Germans planned counterattacks (**e.g., Panzer divisions**), but Hitler’s **delayed orders** prevented an immediate and effective response. |
| **Attackers used cyber warfare for deception** | Cyber warfare **did not exist** in WWII; deception was carried out using **fake radio transmissions and dummy tanks** instead. |
| **Unexpected diplomatic ceasefire mid-battle** | There was **no diplomatic ceasefire** during the Normandy battle; the fighting continued without pause. |
| **Attackers faced major internal command disagreements** | While minor disagreements existed, the **Allied High Command (Eisenhower, Montgomery, Bradley, etc.) remained mostly united** in their strategy. |
| **Defenders received last-minute reinforcements from another front** | German reinforcements were **blocked by the destruction of bridges and roads** due to Allied bombing, preventing a large-scale reinforcement. |

Total False Positives: **5**

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

These are **real events from the Battle of Normandy that the LLM failed to predict**.

| **Event** | **Definition** |
| --- | --- |
| **Allies used specialized engineering units to clear obstacles** | U.S. **Army Rangers and British Commandos** used **demolitions to clear bunkers, minefields, and beach defenses**. |
| **Hedgerows slowed down the Allies’ inland movement** | The Normandy countryside had **thick hedgerows (bocage)** that **slowed tanks and infantry**, making progress inland difficult. |
| **Hitler’s direct intervention delayed German counterattacks** | Hitler personally **refused to release Panzer divisions** early, believing Normandy was a distraction for an attack on Pas-de-Calais. |
| **Allies created artificial harbors (Mulberry Harbors)** | To support supply lines, Allies **built temporary harbors**, allowing unloading of supplies directly onto the beaches. |
| **Allies used underwater pipelines (PLUTO) to transport fuel** | The **Pipeline Under the Ocean (PLUTO)** allowed Allies to **pump fuel from England to France**, ensuring steady supplies. |
| **Rommel attempted but failed to convince Hitler of the real threat** | Field Marshal **Erwin Rommel tried to warn Hitler** that Normandy was the real invasion, but Hitler ignored him. |
| **German command structure suffered from miscommunication** | The **German High Command lacked clear communication**, leading to confusion in executing defensive operations. |
| **Soviet advances were influenced by the timing of D-Day** | The Soviet Union launched **Operation Bagration** on the Eastern Front **shortly after D-Day**, forcing Germany to split its forces. |

Total False Negatives: **8**

***Final Results***

| **Metric** | **Value** |
| --- | --- |
| True Positives (TP) | **19** |
| False Positives (FP) | **5** |
| False Negatives (FN) | **8** |
| Precision | **79.17%** |
| Recall | **70.37%** |
| F1-Score | **74.51%** |

***Analysis and Insights***

* The **LLM performed well in identifying major historical patterns** but missed **some operational details**, especially regarding engineering units and post-landing terrain challenges.
* **Most incorrect predictions (False Positives) were about modern war concepts (cyber warfare, ceasefires, command issues), which do not apply to WWII.**
* **Missed events (False Negatives) included tactical innovations (Mulberry Harbors, PLUTO) and internal German command problems.**
* **Despite this, the LLM provided a highly accurate battle prediction with an F1-score of 74.51%, meaning it successfully captured most key aspects of the Normandy invasion.**