Please provide the following information to be shared with on capstone information exchange platform:

1. **Project number**: S17 - 29

2. **Project title (as will appear on the poster)**: WikiClassify

3. **Team members**: Brian Faure, Wanshu (Wayne) Sun, Nathan Kjer, Luke Wielgus

4. **Adviser(s) name(s)**: Waheed Bajwa

5. **Up to 10 keywords that will help to classify the project:**
   - Machine learning
   - Web development
   - Database management
   - Big data
   - Python
   - C++
   - Ruby on Rails
   - Parsing
   - Text classification
   - Natural language processing
   - Web browser extension

6. **Project abstract (up to 200 words) to be shared with judges:**

   We propose a system that will evaluate and display the quality and importance of every article on the English Wikipedia, with justifications for each classification. Quality and importance will be measured using Wikipedia’s existing categories (featured, A, good, B, C, start, and stub quality; top, high, mid, and low importance). A variety of models to accomplish this task will be compared and selected based on their abilities and performance.

   Wikipedia visitors often encounter articles that are not of sufficient quality. For viewers, identifying problem articles allows them to determine if what they are looking at is accurate. For editors, this information allows them to decide which articles need editing, and for what reasons. Because Wikipedia covers a large portion of the known, defined, English topics; the same methods could potentially be used to identify other problem webpages on the internet and make their classifications known. This could eventually lead to an internet that is intelligently filtered to have content of higher quality, containing less misinformation.