

Applying Aspect-Based Sentiment Analysis to Understand International Fan Reception of Thai Entertainment

Student	George Obinna Oguejiofor
Student ID	george.ogue@kmutt.ac.th
Programme	Bachelor of Science in Computer Science
Advisor	Assoc. Prof. Vajirasak Vanijja
Co-Advisor	Prof. Jonathan H. Chan (Intelligent Systems & Application — AI & Entertainment)
Research Method.	Prof. Debajyoti Pal (Research Methodology)
Submission Date	October 2026

Abstract

Thai drama series, including titles such as *2gether*, *I Told Sunset About You*, and *Girl From Nowhere*, have attracted tens of millions of international viewers across Southeast Asia, East Asia, Europe, and the Americas. Fan review websites such as MyDramaList.com now hold hundreds of thousands of English-language reviews of Thai titles, making them one of the largest untapped sources of data on how international audiences feel about Thai entertainment. Despite this, no published study has used data methods to look at what international fans like or dislike about Thai drama.

The study will draw on a dataset of 1,577 English-language fan reviews covering 20 Thai drama titles across five genres, containing 25,808 opinion pairs labelled across five aspects: Acting, Storyline, Original Soundtrack (OST), Production Quality, and Cast Chemistry. A baseline evaluation of VADER will demonstrate its limitations on this domain: preliminary testing shows it classifies 96.5% of reviews as positive and fails to distinguish how fans feel about different aspects of the same show. To address this, a three-component machine-learning pipeline called GeorgeAspect will be built using scikit-learn. Component A will use TF-IDF with Logistic Regression to classify each review as Positive, Neutral, or Negative (target: 84% accuracy). Component B will use TF-IDF with LinearSVC to produce a separate sentiment score for each of the five aspects (target: above 80% per aspect). Component C will use TF-IDF with Ridge Regression to predict the numeric star rating a reviewer assigns on a scale of 1 to 10 (target: $R^2 \geq 0.45$). All components will be trained and evaluated on an 80/20 train-test split with five-fold cross-validation.

Four tentative research questions guide the study: (1) *What* sentiment patterns exist across the five aspects: Acting, Storyline, OST, Production Quality, and Cast Chemistry, in international fan reviews of Thai dramas, and which aspects matter most to audiences? (2) *Why* do international audiences engage with Thai drama content, and which aspects of the viewing experience (acting, storyline, soundtrack, production quality, or cast chemistry) most influence their overall satisfaction? (3) *How* do fan sentiment patterns across genres reveal which categories of Thai drama attract the broadest and most positive international reception, and what can content creators and streaming platforms learn from those patterns? (4) *Where* do international fan opinions diverge most — across drama genres, across aspects, or between the language of a review and the numeric rating the fan assigns? The project will apply the Agentic Engineering Framework (VC2AE), covering four steps: SPECIFY, HARNESS, VERIFY, and CONTROL. All findings will be documented in a conference-style paper.

Keywords

Aspect-Based Sentiment Analysis (ABSA) · Thai Drama · International Fan Reception · MyDramaList · Opinion Mining · TF-IDF · Logistic Regression · LinearSVC · Ridge Regression · scikit-learn · Cross-Cultural NLP · Entertainment Analytics · GeorgeAspect · Agentic Engineering · VC2AE Framework
