

**Title :** A Tool Using Web Scraping and Text Analysis Techniques for Monitoring Retail Lending Products

**Author(s) :**

1. Benjaporn Deawkanokrath
- 2.
- 3.

**Student ID :** 640510726

**Student ID :**

**Student ID :**

**Major :** Data Science

**Advisor(s) :**

1. Assistant Professor Dr. Phimpaka Taninpong
2. Lecturer Dr. Salinee Thumronglaohapun
- 3.

**Type of presentation\* (choose 1) :**

**Oral Presentation** (เฉพาะ ตัวแทนศ.ที่สาขาเลือกให้นำเสนอแบบบรรยาย)

**Poster** (กรณี นำเสนอผลงานปัญหาพิเศษ/การค้นคว้าอิสระ)

**Cooperative Education** (กรณี นำเสนอผลงานสหกิจศึกษา)

## ABSTRACT

This project aims to develop a tool to enhance efficiency in examining retail credit products by reducing the time required for verification. The research focuses on two primary groups of financial institutions: (1) Domestic systemically important banks (D-SIBs) and (2) Financial institutions with the highest outstanding loan volumes. The verification framework adheres to the regulatory guidelines set by the Bank of Thailand. The development of this tool integrates web scraping techniques and text analysis methodologies, including large language models (LLMs), text extraction, and regular expressions. The tool has been deployed via Streamlit Community Cloud, and its performance has been systematically evaluated through a comparative analysis of the extracted data against actual information available on financial institutions' websites. The results indicate that the tool is capable of accurately summarizing product information in accordance with website content and retrieving all relevant PDF file URLs without omission. This capability significantly improves accessibility to essential information, thereby streamlining the verification process for auditors. However, the tool is not yet equipped to directly assess regulatory compliance, as detailed product information is primarily embedded within PDF documents. Further advancements are required to extend the tool's functionality for accessing and analyzing textual data within PDFs.

\*Type of presentation must be matched with an option you choosing on student upload system.

\*\*The abstract can be more than one page and must be approved by project advisor before upload.