What high school financial education learning objectives are overlooked on a state-by-state basis as compared to national recommendations?

Word Count:5287

Introduction

In recent years, financial education has come to the forefront of state education policy debates. Stakeholders increasingly recognize that equipping students with financial skills is essential for their financial future and a productive national economy (Mandell & Klein, 2009). As of March 2025, twenty-six states mandate students to take a full semester course dedicated to financial education as a high school graduation requirement - a considerable increase from zero states in 2008 (Ramsey Solutions, 2023). However, the momentum to mandate financial education in high schools has outpaced curriculum writers in their efforts to make it effective. For example, Kaiser and Lusardi (2024) find that most financial education programs do not affect long-term financial behavior (Kaiser & Lusardi, 2024). Similarly, Cole and colleagues found that additional math courses had superior effects on students' long-term financial behavior (Cole et al., 2015). Since the efficacy of financial education depends on the curriculum quality and comprehensiveness, the studies above raise concerns regarding the holes in financial education. These gaps in financial education are difficult to identify with manual inspection.

Without comprehensive data on the overlooked learning objectives in education, curriculum writers cannot improve the quality of financial education. This study compares state learning objectives to the Council for Economic Education's national recommendations to identify the gaps in curriculum nationwide. Plugging these overlooked gaps will allow legislators and curriculum writers to improve student education. With a Natural Language Processing (NLP) framework, the researcher can systematically identify the semantic similarity between learning objectives, automating educational standards review (Fanni et al., 2022). The NLP framework paves the way for an objective and efficient approach to improving financial education, so that students nationwide can manage all financial matters in adult life.

Literature Review

American high schools aim to equip students with the required skills to handle adult life. Among these skills, financial literacy stands out as the cornerstone for future success. First defined by The Jump$tart Coalition, financial literacy is the "ability to use knowledge and objective skills to manage one's financial resources effectively for lifetime financial security" (Sherraden, 2010). Based on this definition, effective courses should enhance understanding of financial topics immediately after implementation and guide individuals to grow and protect their assets throughout life. Many researchers have acknowledged that several factors contribute to a program's long-term impact on a student's financial behavior. For example, Wager and colleagues find that financial education programs that are shorter than seven weeks have significantly less impact on students' positive financial behavior months after implementation (Wagner & Walstad, 2018). More importantly, researchers Lusardi and Mitchell found that programs offering a comprehensive curriculum perform significantly better than those focusing on a single topic. Specifically, they found an improvement in essential real-life skills such as regular savings habits, positive investment decisions, and debt avoidance (Lusardi & Mitchell, 2011).

Since new developments in mobile payments, financial technology, identity theft, and philanthropy have changed the nature of financial education, the necessity to create comprehensive curricula has become apparent (Au & Kauffman, 2008). Nonetheless, mandated financial literacy programs throughout the states fail to cover modern financial products and to prepare students for the new complexities of adult life (Heath, 2016). If high schools are to effectively prepare students for developments in real-world finance, then adjusting the curriculum is just the first critical step.

Importance of Financial Literacy Education

Legislators often question the importance of courses exclusively dedicated to financial education, particularly when compared to additional math courses or briefly covering finance in other subjects (Cole et al., 2015). For example, a meta-analysis of educational initiatives in America by the National Bureau of Economic Research (2001) indicated that these programs positively impact future saving rates (Kaiser et al., 2020). In contrast, meta-analysis from the National Endowment for Financial Education showed that financial interventions only explain 0.1% of the variance in financial behavior (Fernandes et al., 2014).

Some researchers attribute this heterogeneity in results to education intensity and delivering financial education at the "teachable moment" (Kaiser & Menkhoff, 2017). Other researchers suggest a more straightforward rationale: financial education, like any other mandated course requirement, cannot function to its full capability without quality implementation.

Other than education itself, several structured studies have found an individuals' financial literacy to be a strong indicator of their financial standing. Lusardi and Tufano found that those with low levels of financial literacy are more likely to have debt problems (Lusardi & Tufano, 2009) Similarly, Lusardi and Mitchell found those with higher levels of financial literacy tend to have more assets, superior debt management, and better quality of life in adulthood (Lusardi & Mitchell, 2011). Therefore, financial literacy education should not be discounted entirely, but revisited and made more comprehensive, so students may acquire literacy and effective financial behavior. Unfortunately there is no existing data into the gaps in financial education. However, if researchers can identify the overlooked topics in financial education then the process to improve curriculum can be replicated nationwide.

The Gap in Financial Education Research

While much of the existing literature has focused on the efficacy of education or the need to improve it, few studies systematically identify the missing learning objectives in states' curricula. Generally, states with a limited budget do not have the resources to continuously update their curriculums to national standards as the process requires a committee of professionals and months of work (*The Curriculum Bridge*, 2025). Thus, while other standards exist, such as those made by the National Financial Educator Council, The Council for Economic Education (CEE) has had the most thorough development and national recognition (Bosshardt & Walstad, 2014). The CEE is a nationally recognized, nonpartisan organization that developed the National Standards for Personal Financial Education. Without a national dataset comparing state learning objectives of financial literacy education with CEE's national recommendations, state financial curriculum writers cannot find the gaps in the curriculum (Kasman et al., 2018). How can financial education be improved if the populace cannot specify the material being taught?

Previously, manual alignment was the norm for improving this curriculum. One notable concern, however, is that this process takes months and is subject to human mistakes and subjective interpretation (Henning & Johnston-Rodriguez, 2018). Manual alignment requires the formation of a committee of professionals. However, educational researchers and policymakers typically have limited time, resulting in limited to no efforts to improve curriculum quality (Fessler et al., 2019).

Natural Language Processing, a subsection of artificial intelligence, offers a scalable, unbiased, and unexplored solution to measuring the nuanced content coverage of state financial education documents (Gruetzemacher, 2022).

Natural Language Processing (NLP) as a Scalable Solution

Researchers have acknowledged the need to improve education without the limitations of manual methods across multiple subjects. Of the automated methodologies for examining educational documents, natural language processing has stood out because of its potential to identify key ideas, learning goals, and depth of coverage for financial literacy topics (Sushree Bibhuprada B. Priyadarshini et al., 2020). Multiple studies demonstrate the capabilities of natural language processing in educational research.

To find gaps in 15 US history textbooks throughout Texas, Lucy et al. used three natural language processing techniques: lexicons, word embeddings, and topic models. They found that Latinx people and women were virtually absent from textbook material, and white men were continuously overrepresented. Additionally, they found word embeddings to be the most effective methodology (Lucy et al., 2020).

Similarly, Gregory Camilli utilized word embeddings to quantify semantic sentences meaning alignment between Common Core State Standards and the 2026 item specifications of the National Assessment of Educational Progress (NAEP) in grade four mathematics (Camilli, 2024). While Camilli's work embedded 5000 educational objectives, the process took less than twenty minutes and used free open-source code with a 90% accuracy rate.

To date, there are no datasets quantifying the overlooked topics in financial education, and there have been no documented attempts by researchers to create one using natural language processing. Logically, since NLP has worked in educational contexts, it can improve financial education. Ultimately, this scalable solution will help this research paper discover the most overlooked learning objectives on a state-by-state basis and nationwide, and this data can improve financial education in classrooms across America.

Method

This method section aims to answer the question, *"What high school financial education learning objectives are overlooked on a state-by-state basis as compared to national recommendations?"* To answer the research question, this study follows the methodologies of Lucy et al. (2020) and Camilli (2024). Specifically, the use of natural language processing word embeddings that quantify semantic similarity between educational content. This project modified OpenAI's semantic similarity guide code and open-source code for this research question, which can be viewed in the appendix for transparency, reproducibility, and future scholarly advancement.

The image below—Figure 1—details the final product workflow for this research paper after many inquiry choices. The progression of these led to the selection and modification of the NLP text embeddings workflow.

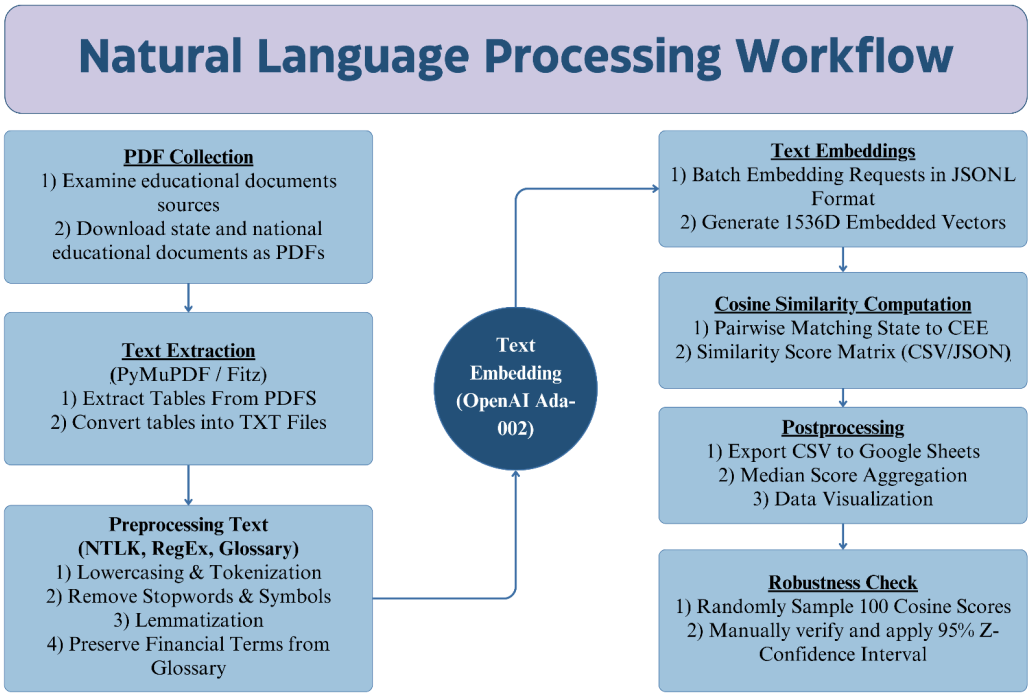

Figure 1 - NLP Workflow for Measuring Alignment Between Learning Objectives

Inquiry Choices That Led to The Natural Language Processing Workflow

The researcher considered multiple methodologies to quantify alignment between state-mandated high school financial literacy learning objectives and national recommendations. Initially, the researcher would manually use the Youth Financial Education Curriculum Review tool, developed by the Consumer Financial Protection Bureau (*Youth Financial Education Curriculum Review Tool | Consumer Financial Protection Bureau*, 2023). However, the tool is intended for committees composed of curriculum designers, policy professionals, and financial literacy experts, and each review session was estimated to take four or more hours per group. After the PDF collection phase, there were 5,000 learning objectives to be inspected, which would be impossible to do in the time frame of this course. It became evident that a more scalable, consistent, and objective method was necessary.

The study first explored Latent Dirichlet Allocation (LDA), a standard topic-modeling algorithm. However, LDA is better suited for finding overlooked (latent) topics rather than comparing alignment between two sets of structured documents (Vayansky & Kumar, 2020).

Then the researcher considered Term Frequency–Inverse Document Frequency (TF-IDF) as it can assign relative importance to terms within a document. Yet, TF-IDF treats words as isolated units and does not understand contextual or semantic relationships between concepts such as “credit risk” versus “credit report” (Christian et al., 2016). This limitation made it unsuitable for high-fidelity alignment between detailed educational objectives. Recognizing these constraints, the researcher adopted a vector-based semantic approach using OpenAI’s text-embedding-ada-002 model. This model converts each learning objective into a 1536-dimensional vector that captures its contextual meaning, allowing for robust and mathematically sound comparisons using cosine similarity.

PDF Collection

The first half of the methodology section focuses on preparing text for the OpenAI model to analyze, as shown in Figure 1 above. The first step in this process is to find relevant, accurate, and organized documents for the OpenAI model. Logically, a curriculum for a full-year course is more comprehensive than a single unit integrated into a math or social studies course. To make this apples-to-apples comparison, the researcher omitted states that do not mandate courses entirely dedicated to financial education (Bernheim et al., 2001). This criterion led to the selection of the following twenty-six states: Alabama, California, Connecticut, Florida, Georgia, Indiana, Iowa, Kansas, Louisiana, Michigan, Minnesota, Mississippi, Missouri, Nebraska, New Hampshire, North Carolina, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Utah, Virginia, West Virginia, Wisconsin.

State PDFs were downloaded from Banzai because documents are up to date and organized in neat rows and columns, which is optimal for simplistic NLP processing (*Life Literacy Courses: Align Education & State Standards*, 2025).

As for the national benchmark, this research paper uses the personal finance standards of education recommended by the Council for Economic Education (CEE) because they are the most comprehensive, nationally recognized, and highlight the importance of weighing the benefits and costs of a decision rather than telling student how to think or operate (Bosshardt & Walstad, 2014). Overall, these standards offer a clear cornerstone for policymakers and educators to evaluate the usefulness of financial education curricula. Moreover, aligning a state's standards with CEE recommendations can increase the possibility that students will get a comprehensive financial education, equipping them with essential skills for long-term financial independence (Olsen, 2017).

Text Extraction and Preprocessing

The second step in natural language processing is to turn extracted PDF documents into machine-readable text. However, not all texts are focused on financial education. For example, introduction pages typically outline how a document should be utilized and by whom the document was developed. Thus, these pages are omitted from comparison. Moreover, there is a crucial difference between standards of education and learning objectives. The CEE focuses on six financial umbrella topics: Earning Income, Spending, Saving, Investing, Managing Credit, and Managing Risk. Each umbrella topic has education standards, which are broad ideas. Specifically, each standard has two to four specific learning objectives focusing on key financial concepts and skills. This research paper only used the Python library PyMuPDF (fitz) and pandas data frame to extract and organize learning objectives (see the appendix for full code documentation). The data frame was organized in a JSON document, a prominent text-based approach to structure data that can be sent to embedding models (Lee et al., 2021).

Following Camilli's (2024) and Lucy et al. (2020) methodologies, the researcher then processed the CSV file through tokenization, lemmatization, and stopword removal (Siino et al., 2024). Typically, by removing punctuation, irrelevant formatting, and other noise, the NLP analysis can efficiently capture the intended context of text. However, tokenization splits terms such as "stock market" into separate words "stock" and "market." To maintain the semantic meaning of these phrases, multi-word financial terms like "credit_score" or "stock_market" are treated as single semantic units with underscores. Therefore, this study differs from Camilli and Lucy et al.'s methods by retaining the meaning of multi-word financial terms. This process reinforces the model's accuracy as the semantic meaning of two individual words is vastly different from the combined meaning of a phrase (Chai, 2022).

Text Embeddings

The purpose of preprocessing text is so that the embedding model can efficiently and accurately create vectors representing each learning objective's meaning. For example, the word "*bank"* could represent the side of a river or a financial institution. Vector embeddings use context words such as "*money*" or "*water*" to identify the topic as either a river or a financial institution (Aiken, 2023). The researcher chose to use OpenAI's embedding model as it understands the nuanced meaning of various topics and avoids the strenuous local model training of alternatives such as the "*BERT*" (Kalyan, 2024). The figure below shows the conversion of the learning objectives as JSON files into embedded vectors.

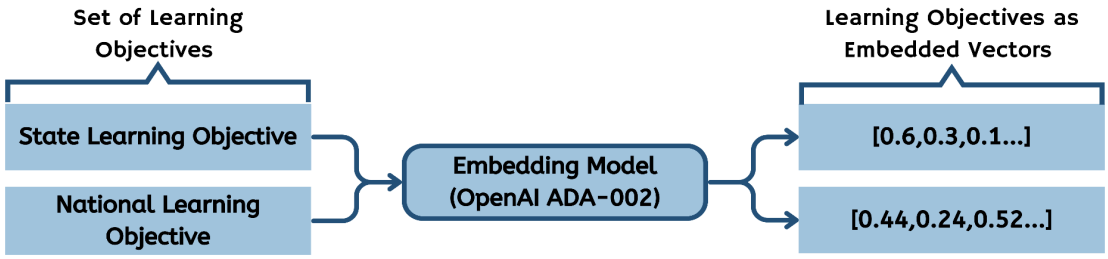

Figure 2 - Learning Objective Embedding Process

One significant difference to Camilli's research is that this paper revises the embedding model from "*text-embedding-3-large"* to "*text-embedding-ada-002."* The researcher chose the latter application programming interface (API) because the large model is for entire textbooks, which would be out of scope for this project, focusing on three to eight-page documents. Then, the embedded vectors are compared with various methods to determine semantic similarity, which, once interpreted, allow for the quantitative representation of how a state's learning objectives align with national standards. This process adequately addresses the gap in the research by providing an updated method to measure alignment in financial education and by creating a dataset that quantifies the overlooked topics in financial education.

Cosine Similarity Computations

After the text-embedding-ada-002 model creates embedded vectors, the next step involves calculating their similarity. While these calculations were made through the Python interface (fully documented in the appendix), this paper represents the computations as mathematical notation for research replication purposes.

Let $S = s_1, s_2, \ldots, s_n$ represent the set of state financial literacy learning objectives, and let $C = \{c_1, c_2, \ldots, c_m\}$ represent the set of national learning objectives from the Council for Economic Education. Each learning objective is embedded as a 1536-dimensional vector, and the figure below shows the transformation of learning objectives into vector quantities through the function F(·).

$$\vec{s}_i = f(s_i), \quad \vec{c}_j = f(c_j) \quad \text{where } \vec{s}_i, \vec{c}_j \in \mathbb{R}^{1536}$$

Figure 3 - Numerical Representation of Embedding Function F(·).

Initially, Euclidean distance was considered for measuring alignment between state and national learning objectives. However, after initial trials, it became apparent that this metric exaggerated dissimilarity because the CEE document was considerably larger than the state curriculum documents. Thus, cosine similarity scores were utilized to quantify alignment between state $\vec{s}_i$ and national $\vec{c}_j$ learning objective vectors as they are more accurate for disparate document lengths (Qian et al., 2004). The figure below illustrates the vast disparity between Cosine Similarity and Euclidean Distance and the cosine formula.

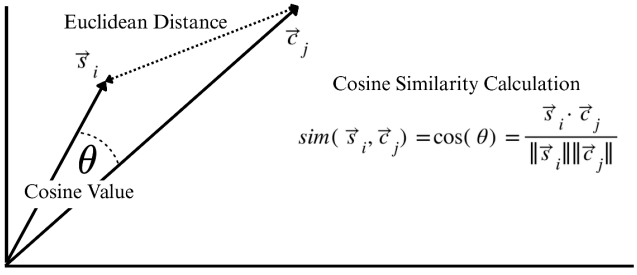

Figure 4 - Numerical Representation of Embedding Function F(·)

Post Processing

The post processing phase is essential in order to understand the cosine similarity scores created by the Python algorithm. After the embedding model matches each pair of cosine similarity scores, the Python script function *scipy.spatial.distance.cosine* calculates cosine similarity (Gunawan et al., 2018). Each national learning objective was matched to the highest cosine similarity counterpart in the state document, even if duplicates occurred. Then, using the function *df.to_csv(cosine_similarity_scores.csv),* the data frame of cosine similarity scores can be converted into a comma-separated values (CSV) file. This file was imported into Google Sheets with 5337 rows. Two of the rows out of 5337 are shown for reference.

Cosine Similarity	National Standard	State Standard
1	*12-1a. Describe how credit card grace periods, methods of interest calculation, and fees affect borrowing costs.*	*a. Explain how credit card grace periods, methods of interest calculation, and fees affect borrowing costs.*
0.25	*12-1b. Compare the cost of borrowing $1,000 using consumer credit options that differ in rates and fees.*	*a. Explain how credit card grace periods, methods of interest calculation, and fees affect borrowing costs.*

Figure 5 - Exported CSV file

The researcher organized the cosine similarity scores with the google sheets function *=INDEX(data)* to create two tables: the first focusing on state-by-state analysis and the second focusing on nationwide analysis. For the nationwide analysis, the function *=MEDIAN(data)* found the median cosine similarity score for a single learning objective across all states (Osborne & Overbay, 2004). The researcher used the median instead of the mean (average) because the mean can be distorted by outliers. Thus, the median is a more accurate representation of the population. Then, functions *=SMALL(data, n)* allowed the researcher to find the smallest cosine similarity scores, which represent the most overlooked learning objectives.

Finally, the researcher created twenty-seven separate graphs. The first twenty-six graphs showed the five lowest cosine similarity scores for each of the twenty-six states. Most importantly, the twenty-seventh graph showcased the five lowest cosine similarity scores nationwide. Overall, these graphs answered the overarching research question.

Robustness Check

To estimate the proportion of correct cosine similarity scores, the researcher used a one-sample z-interval for proportions. Z-intervals are the industry standard in Advanced Placement Statistics textbooks and data analysis (Singh & Masuku, 2014). To estimate the accuracy rate of the cosine similarity scores, the researcher randomly selected a sample of 100 cosine similarity scores with the TI-84 calculator's random integer function RandInt (1, 5357, 100). This 100-sample subset is selected to capture a representative mix of financial topics (e.g., saving, credit, investing), varied linguistic complexity, and standards from multiple states. All data was moved to a separate Google sheet document, and each of the 100 cosine similarity scores was then examined for authenticity. Then, cosine similarity scores were inspected, with accurate scores receiving the value of 1 and inaccurate scores receiving the value of 0. The figure below shows this process.

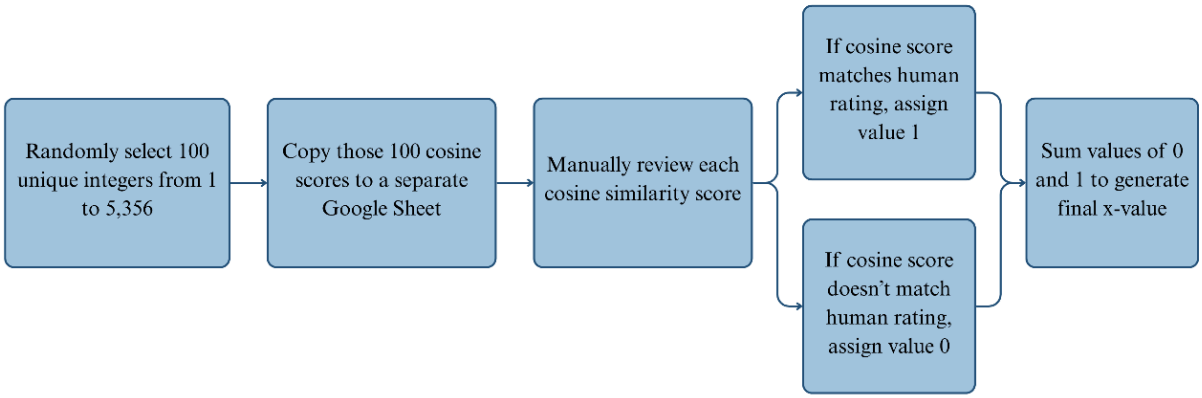

Figure 6 - Estimating Accuracy of Cosine Similarity Scores with Z-Interval

The researcher followed industry standards using the TI-84 calculator to input the *x* and *n* values to compute the confidence interval, which estimates the proportion of accurate similarity ratings in the entire data set of 5357 learning objectives. This manual validation step saves time as each of the 5357 learning objectives does not need to be inspected while supporting credibility, which is crucial for policy recommendations in the implications section.

Results

The most crucial finding in this research paper is the accuracy of the natural language processing model. This model compared semantic similarity between 5357 learning objectives - a task that, twenty years ago, would have required a committee of teachers and financial professionals because of the nuance in each term. Across the 100 randomly sampled cosine similarity scores, 83 of them were deemed accurate by the researcher. When implemented into the Z-proportion, this statistic created an interval of [0.756, 0.903]. Thus, this paper has statistical evidence that with 95% confidence, the proportion of accurate cosine similarity ratings is between 75% and 90% (O'Brien & Yi, 2016).

The high accuracy of the model showcases the tool's efficacy in improving financial education. The simple random sample (n=100) of all learning objectives is listed in Appendix D. As shown in Appendix A or Figure 7, the five most overlooked learning objectives all have cosine similarity scores near zero (the full list of data is analyzed in Appendices A-C). These learning objectives had median cosine scores across all states lower than 0.2, indicating a minimal alignment to national financial education benchmarks.

Median Cosine Similarity Score	**Top Five Overlooked Learning Objectives Nationwide**
0	12-6c. Define key rental contract terminology, including lease term, security deposit, grace period, and eviction.
0.09	12-4c. Discuss possible explanations for the persistence of race and gender pay gaps.
0.14	12-11b. Investigate the effects of bankruptcy on assets, employment, and future access to credit.
0.16	12-2a. Select a preferred location for a savings account based on comparison of interest rates and fees at different types of financial institutions.
0.16	12-7b. Compare the expense ratios for several mutual funds.

Figure 7 - The Five Most Overlooked Learning Objectives Nationwide

Overlooked Learning Objectives on a Category-by-Category Basis

The twenty-six states in our dataset had nuanced disparities, but this diagram focuses on the general trends to conduct a state-by-state analysis. First, the researcher separated all learning objectives into the six broad categories listed by the CEE. Then, the semantic similarity of each category from zero to one, compared to the CEE document, is visually represented on a color graph, with lighter shades of blue indicating lower similarity and darker shades indicating higher similarity. Figure 7 shows the color gradings of each state in the six main financial education categories.

State	Earning Income	Spending	Saving	Investing	Managing Credit	Managing Risk
Alabama	0.28	0.52	0.37	0.52	0.25	0.41
California	0.93	1	1	0.31	1	1
Connecticut	1	1	1	0.33	1	1
Florida	0.4	0.65	0.49	0.41	0.44	0.525
Georgia	0.49	0.43	0.32	0.45	0.43	0.365
Indiana	0.485	0.49	0.3	0.38	0.36	0.38
Iowa	0.425	0.5	0.4	0.29	0.3	0.49
Kansas	1	1	1	0.33	1	1
Louisiana	0.57	0.57	0.3	0.59	0.435	0.43
Michigan	0.48	0.47	0.44	0.52	0.55	0.41
Minnesota	0.655	0.52	0.34	0.53	0.34	0.435
Mississippi	0.275	0.35	0.36	0.35	0.49	0.315
Missouri	0.53	0.49	0.45	0.41	0.38	0.495
Nebraska	0.585	0.56	0.42	0.6	0.46	0.415
New Hampshire	0.39	0.58	0.25	0.44	0.41	0.425
North Carolina	0.54	0.38	0.43	0.33	0.48	1
Ohio	0.615	0.47	0.34	0.43	0.57	0.46
Oregon	0.51	0.57	0.38	0.58	0.55	0.395
Pennsylvania	0.5	0.55	0.48	0.5	0.37	0.37
Rhode Island	1	1	1	0.33	1	1
South Carolina	0.36	0.42	0.5	0.39	0.29	0.555
Tennessee	0.555	0.53	0.42	0.48	0.37	0.43
Utah	0.595	0.23	0.35	0.42	0.45	0.445
Virginia	0.64	0.27	0.49	0.28	0.29	0.535
West Virginia	0.405	0.62	0.42	0.33	0.46	0.36
Wisconsin	0.485	0.5	0.43	0.51	0.57	0.295

Figure 8 - Learning Objectives Separated into Six Categories

This graph shows intense dark squares in California, Connecticut, Kansas, and Rhode Island, which are close partners with the CEE. Nonetheless, investing learning objectives are overlooked not only in most states but even in those directly partnering with the CEE. These results suggest that significant improvements should be made to improve the curriculum, as the general trend showed scores of around 0.5. Most notably, investing is the least represented topic across the board in financial education, while earning income is stronger. These results suggest a redistribution of resources to improve curriculum comprehensiveness.

Overlooked Learning Objectives on a State-by-State Basis

The most valuable results of this paper involve the clear classification of the overlooked learning objectives on a state-by-state basis. The first key observation is that states Alabama, Indiana, and Michigan all have cosine similarity scores of zero for the learning objective "Complete IRS Form W-4." This finding raises the question regarding the discussion of taxes in classes. Another key trend existed in states Minnesota, Missouri, and Tennessee; these had similarity scores of zero for "explain the role of federal regulators in financial markets." The absence of knowledge regarding the SEC, FINRA, or CFPB reduces investors confidence in investment vehicles and overall market participation (Pan, 2012). Thus, the lack of these learning objectives not only hurts the students but also the market economy. Lastly, the absence of coverage regarding mutual funds, CDs, the S&P 500, and loan agreements across states aligns with general findings in the category-by-category analysis, showing investment learning objectives to be underrepresented across the dataset.

State	Overlooked Learning Objective (Cosine Score = 0)
Alabama	12-9a. Complete IRS Form W-4.
California	12-7a. Investigate the federal and state tax rates applicable to different sources of income.
Connecticut	12-7b. Compare sales tax rates paid on different types of goods in their state and for online purchases.
Florida	12-10d. Report the average benefit paid to a retiree living on Social Security today.
Georgia	12-8b. Describe the advantages of investing through a tax-deferred account such as an IRA or 401(k) versus a taxable account.
Indiana	12-9a. Complete IRS Form W-4.
Iowa	12-10d. Report the average benefit paid to a retiree living on Social Security today.
Kansas	12-1c. Assess their personal risk tolerance using an online tool or worksheet.
Louisiana	12-4a. Describe the impact of inflation on prices over time.
Michigan	12-9a. Complete IRS Form W-4.
Minnesota	12-12a. Explain the role of federal regulators in financial markets.
Mississippi	12-7b. Compare the expense ratios for several mutual funds.
Missouri	12-12a. Explain the role of federal regulators in financial markets.
Nebraska	12-7a. Explain how an employer match of employee contributions to its retirement plan provides an incentive for employees to save.
New Hampshire	12-11a. Evaluate the benefits and costs of gig employment, such as driving for a cab or delivery service.
North Carolina	12-9a. Explain how external influences (e.g. peers, family, or social media) can impact personal savings decisions.
Ohio	12-4a. Describe the impact of inflation on prices over time.
Oregon	12-6a. Identify examples of loans that may require down payments.
Pennsylvania	12-1a. Compare the features of regular savings accounts, money market accounts, and CDs.
Rhode Island	12-1c. Assess their personal risk tolerance using an online tool or worksheet.
South Carolina	12-13a. Explain why investors often compare portfolio performance to a benchmark such as the S&P 500 Index.
Tennessee	12-12a. Explain the role of federal regulators in financial markets.
Utah	12-7b. Compare the expense ratios for several mutual funds.
Virginia	12-8b. Discuss the benefits and costs of purchasing life insurance on the primary earners in a household.
West Virginia	12-3a. Explain why homeownersâ€™ insurance is required by a lender when a homeowner takes out a mortgage.
Wisconsin	12-10d. Compare the costs and benefits associated with for-profit versus non-profit credit counseling services.

Figure 9 - Overlooked Learning Objectives State-by-State

Limitations

The first pertinent limitation was the pitfalls of semantic similarity. Some text embedding models yield cosine high similarity scores because of identical vocabulary. For example, “I like rainy days because they make me feel relaxed” and “I do not like rainy days because they do not make me feel calm” return a cosine similarity score of 0.993 using the API model distilbert-base-uncased (Face, 2021). The OpenAI model may not be able to capture the negative or affirmative nature of sentences. Thus, two learning objectives can be closely related in the vector space yet opposites regarding meaning. Nonetheless, in this research paper, opinions on positive financial behavior are rarely opposite across documents. Thus, semantic understanding, just like human error, is a strong limitation to the paper's findings.

The general-purpose NLP model embedding-ada-002 by OpenAI is to be domain-adaptable (Nussbaum et al., 2024). However, the accuracy of this model can degrade when focused on a specialized field such as financial education. Consequently, the researcher implemented a domain-specific text file of financial terms during the first preprocessing portion of the method section. By connecting financial terms with underscores, the researcher retained the meaning of phrases such as “stock_market.” The general-purpose embedding model embedding-ada-002 is not as effective as a finance-specific benchmark. Specifically, through regression analysis, Tang & Yang (2024) found that the Finance Massive Text Embedding Benchmark (FinMTEB) had an error rate that was 10% less than standard ChatGPT mini embedding software (Tang & Yang, 2024). The general-purpose embedding model may have overlooked more domain-specific topics in financial education, creating a less accurate cosine similarity score and selecting overlooked and overrepresented topics in financial education. Thus, to fix this limitation, researchers should adopt a more domain-specific model.

Input Length Sensitivity and Interval Conditions

The input length sensitivity of the embedding model was a strong limitation to the cosine similarity score accuracy interval of 75% to 90%. Embedding models are designed to handle variable input lengths from multi-page documents to a single financial learning objective, which typically is a short phrase or sentence. In the preprocessing for natural language processing, the researcher used tokenization, stopword removal, and lemmatization (Smelyakov et al., 2020). Specifically, stopword removal or the removal of words such as "the," "a," and "is" may have been an unnecessary preprocessing step. This could have resulted in the discarding of critical information in the semantic meaning of a phrase (Kaur & Buttar, 2018). This process is typically done for entire textbooks, not single sentences, so it was unnecessary for this paper.

The last important limitation to mention is that effective proportion estimation with statistical analysis satisfies three key conditions: there being a simple random sample, each unit being independent from another unit, and the population data having approximate normality (Singh & Masuku, 2014). The cosine similarity scores, when plotted, are right-skewed, with a visible peak near the 1.0 mark, indicating a large cluster of perfectly aligned standards. Therefore, the interval could indicate an inflated proportion of correct similarity scores.

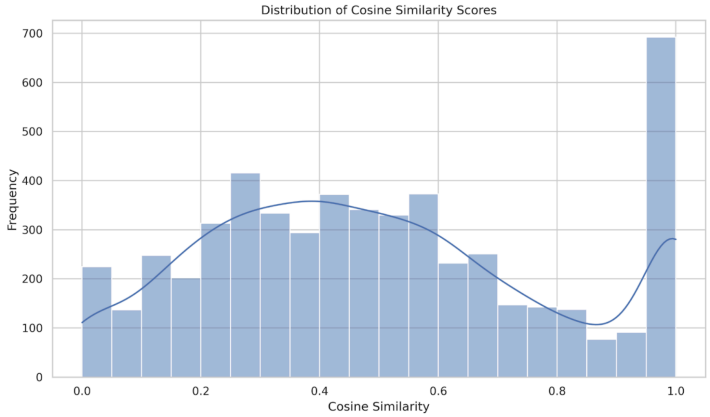

Figure 10 - Cosine Similarity Score Plotted with Strong Visible Peak

Implications

The previous sections, supported by Appendices A-C, all have substantial implications for artificial intelligence research, policy, and curriculum development. In contrast, Appendices E-F document the model for reproducibility and Appendices G-H represent my use of findings.

Understanding where the curriculum lacks allows individuals to improve financial education and, as a result, financial behavior. Overall, this research paper discovered what learning objectives were overlooked state-by-state, with 75% to 95% accuracy. The larger understanding presented is that we don't just need more financial education - we need higher quality financial education.

This study has notable value to the body of AI model literature, expanding on NLP's relevance to how we optimize education research. The critical implication is that this paper validates the use of the text-embedding-ada-002 model in large corpora text comparisons, expanding on the work of Lucy and colleagues who showed that work embeddings can be used to assess semantic coverage in historical context (Lucy et al., 2020). Moreover, the results validate the concerns of Tang and Yang as they established that general-purpose models encounter occasional errors (Tang & Yang, 2024). They suggest the use of Finance Massive Text Embedding Benchmark (FinMTEB) as its accuracy outperforms general models when comparing similar vocabulary (avoid credit vs establish credit). Because this model experienced occasional errors with similar vocabulary, it can be established that AI tools should be used with transparency, as they are not completely accurate. These findings support the general push to a more responsible and ethical interpretation of policy-relevant machine learning findings. Most importantly, this research provides a scalable pipeline for financial education research and a modern approach to financial education policy.

Policy Implications

Another facet implication for this project is to move away from the method and to look towards the dataset itself. By systematically identifying bills in progress with the use of the Next Generation Personal Finance "Bill Tracker," the researcher - or any other individual - can sort by active bills across all states (Next Gen Personal Finance, 2025). Then, the researcher will identify the right audience for writing a letter to support or oppose legislation, which can be found at OpenStates.org. Additionally, state treasurers often have an influential voice in financial education and will be considered for emailing as well (National Association of State Treasurers, 2025).

Following this logic, the researcher shared research findings to representatives Darly Campbell and Dan Daley. They are advocating for a new HB 737 bill in Florida which could add "career readiness and financial education" enhancements to grades 9-12 (The Florida Senate, 2025). This paper proposes to explicitly mention the overlooked learning objectives in Appendices A-C in HB 737. For example, Florida is one of the many states that overlooked key lease agreements and bankruptcy learning objectives. Most college students participate in a dorm agreement or lease agreement, and the absence of education in this process can lead students to be unaware of possible hidden fees or rising interest rates (Sotomayor et al., 2022). The lack of education on lease agreements and bankruptcy means students can face eviction because they failed to understand the grey areas in lease contracts (Cunha, 2025). Thus, a key policy change in Florida would be to explicitly mention bankruptcy and lease agreement topics as a topic in high school classrooms, so students can avoid hidden fees that can lead to bankruptcy (Furth-Matzkin, 2017). Similar to how US History textbooks were updated for the COVID-19 pandemic, the constantly updating nature of financial instruments and practices requires a modern curriculum.

Systemic Inequality and Future Research

Curriculum content plays a major role in students' perceptions of the larger social and cultural historical narrative in America (Sjölund Åhsberg, C, 2024). The absence of coverage regarding key parts of American society will affect how students see and behave in the world, so overlooking key parts of curriculum limits an individual's understanding of their financial history. In the paper by Lucy and colleagues, Women and Latinx groups were repeatedly overlooked in their dataset of 15 Texas US History textbooks (Lucy et al., 2020). Similar to Lucy's findings, this study shows that the learning objective "discuss possible explanations for the persistence of race and gender pay gaps" is the second most overlooked across the entire dataset. Historically, communities of color and lower income households have been the subjects of predatory lending, redlining, and low credit access (Nier & St. Cyr, 2010). Therefore, the lack of discussion regarding this key part of American History is concerning as they have been systematically excluded from wealth-building opportunities for decades (Sjölund Åhsberg, C, 2024). This paper's appendices demonstrate the empirical evidence (A-D), technical implementation (E-F), and active impact (G-H) of natural language processing.

The learning objectives in schools must address the populations that would most benefit from such literacy, and investing can be the key for most people to open the door for financial freedom. The topic can help those with socioeconomic struggles to develop generational wealth that can decrease reliance on public wealthfare systems, helping the American economy holistically. Still, investment learning objectives are overlooked disproportionately in comparison to other subject areas. We don't need more financial education, we need the right financial education. Thus, it is imperative that researchers investigate natural language processing curriculum review models to make our curriculum as comprehensive as possible.

Bibliography

Aiken. (2023, August 19). *Exploring Cultural Differences in Language Meaning: The Intriguing World of Semantics – TCL Lab*. TCL Lab. https://tcllab.org/exploring-cultural-differences-in-language-meaning-the-intriguing-world-of-semantics/

Artifex. (2019). *Tutorial - PyMuPDF 1.24.14 documentation*. Readthedocs.io. https://pymupdf.readthedocs.io/en/latest/tutorial.html

Au, Y. A., & Kauffman, R. J. (2008). The economics of mobile payments: Understanding stakeholder issues for an emerging financial technology application. *Electronic Commerce Research and Applications*, *7*(2), 141–164. https://doi.org/10.1016/j.elerap.2006.12.004

Bernheim, B. Douglas., Garrett, D. M., & Maki, D. M. (2001). Education and saving: *Journal of Public Economics*, *80*(3), 435–465. https://doi.org/10.1016/s0047-2727(00)00120-1

Bosshardt, W., & Walstad, W. B. (2014). National Standards for Financial Literacy: Rationale and Content. *The Journal of Economic Education*, *45*(1), 63–70. https://doi.org/10.1080/00220485.2014.859963

Camilli, G. (2024). *An NLP Crosswalk Between the Common Core State Standards and NAEP Item Specifications*. ArXiv.org. https://arxiv.org/abs/2405.17284

Chai, C. P. (2022). Comparison of text preprocessing methods. *Natural Language Engineering*, *29*(3), 1–45. https://doi.org/10.1017/s1351324922000213

Christian, H., Agus, M. P., & Suhartono, D. (2016). Single Document Automatic Text Summarization using Term Frequency-Inverse Document Frequency (TF-IDF).

ComTech: Computer, Mathematics and Engineering Applications, *7*(4), 285. https://doi.org/10.21512/comtech.v7i4.3746

Cole, S., Paulson, A., & Shastry, G. K. (2015). High School Curriculum and Financial Outcomes: The Impact of Mandated Personal Finance and Mathematics Courses. *Journal of Human Resources*, *51*(3), 656–698. https://doi.org/10.3368/jhr.51.3.0113-5410r1

Collins, E. R. (2024, February 18). *5 Best Ways to Convert Pandas DataFrame GroupBy to List – Be on the Right Side of Change*. Finxter.com. https://blog.finxter.com/5-best-ways-to-convert-pandas-dataframe-groupby-to-list/

Cunha, C. D. (2025). *Living on the Edge College Students and Landlords: the Perceptions, the Misconceptions and the Reality*. SURFACE at Syracuse University. https://surface.syr.edu/honors_capstone/812/

Face, H. (2021, September 29). *Sentence similarity models not capturing opposite sentences*. Hugging Face Forums. https://discuss.huggingface.co/t/sentence-similarity-models-not-capturing-opposite-sentences/10388

Fanni, R., Steinkogler, V. E., Zampedri, G., & Pierson, J. (2022). Enhancing human agency through redress in Artificial Intelligence Systems. *AI & SOCIETY*, *38*. https://doi.org/10.1007/s00146-022-01454-7

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, *60*(8), 1861–1883. https://doi.org/10.1287/mnsc.2013.1849

Fessler, P., Silgoner, M., & Weber, R. (2019). Financial knowledge, attitude and behavior: evidence from the Austrian Survey of Financial Literacy. *Empirica*. https://doi.org/10.1007/s10663-019-09465-2

Furth-Matzkin, M. (2017). On the Unexpected Use of Unenforceable Contract Terms: Evidence from the Residential Rental Market. *Journal of Legal Analysis*, *9*(1), 1–49. https://doi.org/10.1093/jla/lax002

Ghosh, S. (2024, September 23). *A Deep Dive into OpenAI's text-embedding-ada-002: The Power of Semantic Understanding*. Medium. https://medium.com/@siladityaghosh/a-deep-dive-into-openais-text-embedding-ada-002-the-power-of-semantic-understanding-7072c0386f83

Gruetzemacher, R. (2022). *The Power of Natural Language Processing*. Harvard Business Review. https://hbr.org/2022/04/the-power-of-natural-language-processing

Gunawan, D., Sembiring, C. A., & Budiman, M. A. (2018). The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents. *Journal of Physics: Conference Series*, *978*, 012120. https://doi.org/10.1088/1742-6596/978/1/012120

Heath, J. A. (2016). Youth Financial Literacy in the United States: A Patchwork Approach. *Springer EBooks*, 369–380. https://doi.org/10.1007/978-981-10-0360-8_25

Henning, M. B., & Johnston-Rodriguez, S. (2018). Evaluating financial literacy curriculum for young adults with special needs: A review of content, universal design for learning, and culturally responsive curriculum principles. *Citizenship, Social and Economics Education*, *17*(2), 118–135. https://doi.org/10.1177/2047173418789593

Kaiser, T., & Lusardi, A. (2024). Financial Literacy and Financial Education: An Overview. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4803857

Kaiser, T., Lusardi, A., Menkhoff, L., & Urban, C. (2020). Financial Education Affects Financial Knowledge and Downstream Behaviors. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3588131

Kaiser, T., & Menkhoff, L. (2017, May 1). *Does Financial Education Impact Financial Literacy and Financial Behavior, and If so, When?* Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2753510

Kalyan, K. S. (2024). A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Natural Language Processing Journal*, *6*, 100048. https://doi.org/10.1016/j.nlp.2023.100048

Kasman, M., Heuberger, B., & Hammond, R. (2018). *E C O N O M I C S T U D I E S AT B R O O K I N G S A Review of Large-Scale Youth Financial Literacy Education Policies and Programs BROOKINGS INSTITUTION BROOKINGS INSTITUTION i /// A Review of Large-scale Youth Financial Literacy Education Policies and Programs E C O N O M I C S T U D I E S AT B R O O K I N G S.* https://www.brookings.edu/wp-content/uploads/2018/10/ES_20181001_Financial-Literacy-Review.pdf

Kaur, J., & Buttar, P. K. (2018). *A Systematic Review on Stopword Removal Algorithms*. *4*(4), 207–210. https://www.researchgate.net/publication/354950829_A_Systematic_Review_on_Stopword_Removal_Algorithms

Lee, J., Anjos, E., & Satti, S. R. (2021). SJSON: A succinct representation for JSON documents. *Information Systems*, *97*, 101686. https://doi.org/10.1016/j.is.2020.101686

Liang, D. (2024, July 30). *Intro — Getting Started with Text Embeddings: Using BERT*. Medium. https://medium.com/@davidlfliang/intro-getting-started-with-text-embeddings-using-bert-9f8c3b98dee6

Life Literacy Courses: Align Education & State Standards. (2025). Banzai.org. https://banzai.org/curriculum

Lucy, L., Demszky, D., Bromley, P., & Jurafsky, D. (2020). Content Analysis of Textbooks via Natural Language Processing: Findings on Gender, Race, and Ethnicity in Texas U.S. History Textbooks. *AERA Open*, *6*(3), 233285842094031. https://doi.org/10.1177/2332858420940312

Lusardi, A., & Mitchell, O. S. (2011, June 2). *Financial Literacy around the World: An Overview*. Www.nber.org. https://www.nber.org/papers/w17107

Lusardi, A., & Tufano, P. (2009, March). *Debt Literacy, Financial Experiences, and Overindebtedness*. ResearchGate; unknown. https://www.researchgate.net/publication/48264008_Debt_Literacy_Financial_Experiences_and_Overindebtedness

Mandell, L., & Klein, L. (2009). *The Impact of Financial Literacy Education on Subsequent Financial Behavior*. https://files.eric.ed.gov/fulltext/EJ859556.pdf

Mariya Mansurova. (2024, February 13). *Text Embeddings: Comprehensive Guide - TDS Archive - Medium*. Medium; TDS Archive. https://medium.com/data-science/text-embeddings-comprehensive-guide-afd97fce8fb5

National Association of State Treasurers. (2025, April). *Find Your State Treasurer*. National Association of State Treasurers (NAST). https://nast.org/find-your-state-treasurer/

Next Gen Personal Finance. (2025). *2023 Financial Education Bill Tracker*. Www.ngpf.org. https://www.ngpf.org/bill-tracker/

Nier, C. L. I., & St. Cyr, M. R. (2010). A Racial Financial Crisis: Rethinking the Theory of Reverse Redlining to Combat Predatory Lending under the Fair Housing Act. *Temple Law Review*, *83*, 941. https://heinonline.org/HOL/LandingPage?handle=hein.journals/temple83&div=35&id=&page=

Nussbaum, Z., Morris, J. X., Duderstadt, B., & Mulyar, A. (2024, February 2). *Nomic Embed: Training a Reproducible Long Context Text Embedder*. ArXiv.org. https://doi.org/10.48550/arXiv.2402.01613

O'Brien, S. F., & Yi, Q. L. (2016). How Do I Interpret a Confidence Interval? *Transfusion*, *56*(7), 1680–1683. https://doi.org/10.1111/trf.13635

Olsen, S. (2017). *The Implementation of Financial Literacy Education in Public High School Classrooms*. Scholar Commons. https://scholarcommons.sc.edu/senior_theses/172/

OpenAI. (2024). *OpenAI Platform*. Openai.com. https://platform.openai.com/docs/overview

Osborne, J., & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). *Practical Assessment, Research, and Evaluation*, *9*(6), 6. https://doi.org/10.7275/qf69-7k43

Pan, E. J. (2012). Understanding Financial Regulation. *Utah Law Review*, *2012*, 1897. https://heinonline.org/HOL/LandingPage?handle=hein.journals/utahlr2012&div=50&id=&page=

Qian, G., Sural, S., Gu, Y., & Pramanik, S. (2004). Similarity between Euclidean and cosine angle distance for nearest neighbor queries. *Proceedings of the 2004 ACM Symposium on Applied Computing - SAC '04*. https://doi.org/10.1145/967900.968151

Ramsey Solutions. (2023, April 3). *Which States Require Financial Literacy for High School Students?* Ramsey Solutions. https://www.ramseysolutions.com/financial-literacy/states-require-financial-literacy-in-high-school#cookie-banner

Sherraden, M. S. (2010). *Financial Capability: What Is It, and How Can It Be Created?* https://doi.org/10.7936/k7sx6cqx

Siino, M., Tinnirello, I., & Marco La Cascia. (2024). Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers. *Information Systems*, *121*, 102342–102342. https://doi.org/10.1016/j.is.2023.102342

Singh, A. S., & Masuku, M. (2014, November). *Sampling techniques and determination of sample size in applied statistics research: An overview*. ResearchGate. https://www.researchgate.net/publication/341552596_Sampling_Techniques_and_Determination_of_Sample_Size_in_Applied_Statistics_Research_An_Overview

Sjölund Åhsberg, C. (2024). Students' views of historical significance – a narrative literature review - UCL Discovery. *Ucl.ac.uk*. https://discovery.ucl.ac.uk/id/eprint/10191870/1/herj-3143-%C3%A5hsberg.pdf

Smelyakov, K., Karachevtsev, D., Kulemza, D., Samoilenko, Y., Patlan, O., & Chupryna, A. (2020, October 1). *Effectiveness of Preprocessing Algorithms for Natural Language*

Processing Applications. IEEE Xplore. https://doi.org/10.1109/PICST51311.2020.9467919

Sotomayor, L., Tarhan, D., Vieta, M., McCartney, S., & Mas, A. (2022). When students are house-poor: Urban universities, student marginality, and the hidden curriculum of student housing. *Cities*, *124*(103572). https://doi.org/10.1016/j.cities.2022.103572

Sushree Bibhuprada B. Priyadarshini, Amiya Bhusan Bagjadab, & Mishra, B. K. (2020). A Brief Overview of Natural Language Processing and Artificial Intelligence. *Apple Academic Press EBooks*, 211–224. https://doi.org/10.1201/9780367808495-8

Tang, Y., & Yang, Y. (2024). *Do We Need Domain-Specific Embedding Models? An Empirical Investigation*. ArXiv.org. https://arxiv.org/abs/2409.18511

The Curriculum Bridge. (2025). Google Books. https://books.google.com/books?hl=en&lr=&id=pNVr7FL7URsC&oi=fnd&pg=PP1&dq=.+Generally

The Florida Senate. (2025). *House Bill 737 (2025) - The Florida Senate*. Flsenate.gov. https://www.flsenate.gov/Session/Bill/2025/737

Vayansky, I., & Kumar, S. A. P. (2020). A review of topic modeling methods. *Information Systems*, *94*, 101582. https://doi.org/10.1016/j.is.2020.101582

Wagner, J., & Walstad, W. B. (2018). The Effects of Financial Education on Short-Term and Long-Term Financial Behaviors. *Journal of Consumer Affairs*, *53*(1), 234–259. https://doi.org/10.1111/joca.12210

Youth Financial Education Curriculum Review Tool | Consumer Financial Protection Bureau. (2023, August 8). Consumer Financial Protection Bureau.

https://www.consumerfinance.gov/data-research/research-reports/youth-financial-education-curriculum-review-tool/

Appendix

Appendix A - 5 Most Overlooked Learning Objectives Nationwide

0	**12-6c. Define key rental contract terminology, including lease term, security deposit, grace period, and eviction.**
0.09	**12-4c. Discuss possible explanations for the persistence of race and gender pay gaps.**
0.14	**12-11b. Investigate the effects of bankruptcy on assets, employment, and future access to credit.**
0.16	**12-2a. Select a preferred location for a savings account based on comparison of interest rates and fees at different types of financial institutions.**
0.16	**12-7b. Compare the expense ratios for several mutual funds.**

Appendix B - State-by-State Overlooked Learning Objectives

This analysis focuses on learning objectives not previously mentioned in Appendix C. Additionally, this is not a full list but rather a random selection of learning objectives that have a cosine similarity score equivalent to zero.

State	Overlooked Learning Objective (Cosine Score = 0)
Alabama	12-9a. Complete IRS Form W-4.
California	12-7a. Investigate the federal and state tax rates applicable to different sources of income.
Connecticut	12-7b. Compare sales tax rates paid on different types of goods in their state and for online purchases.
Florida	12-10d. Report the average benefit paid to a retiree living on Social Security today.
Georgia	12-8b. Describe the advantages of investing through a tax-deferred account such as an IRA or 401(k) versus a taxable account.
Indiana	12-9a. Complete IRS Form W-4.
Iowa	12-10d. Report the average benefit paid to a retiree living on Social Security today.
Kansas	12-1c. Assess their personal risk tolerance using an online tool or worksheet.
Louisiana	12-4a. Describe the impact of inflation on prices over time.
Michigan	12-9a. Complete IRS Form W-4.
Minnesota	12-12a. Explain the role of federal regulators in financial markets.
Mississippi	12-7b. Compare the expense ratios for several mutual funds.
Missouri	12-12a. Explain the role of federal regulators in financial markets.
Nebraska	12-7a. Explain how an employer match of employee contributions to its retirement plan provides an incentive for employees to save.
New Hampshire	12-11a. Evaluate the benefits and costs of gig employment, such as driving for a cab or delivery service.
North Carolina	12-9a. Explain how external influences (e.g. peers, family, or social media) can impact personal savings decisions.
Ohio	12-4a. Describe the impact of inflation on prices over time.
Oregon	12-6a. Identify examples of loans that may require down payments.
Pennsylvania	12-1a. Compare the features of regular savings accounts, money market accounts, and CDs.
Rhode Island	12-1c. Assess their personal risk tolerance using an online tool or worksheet.
South Carolina	12-13a. Explain why investors often compare portfolio performance to a benchmark such as the S&P 500 Index.
Tennessee	12-12a. Explain the role of federal regulators in financial markets.
Utah	12-7b. Compare the expense ratios for several mutual funds.
Virginia	12-8b. Discuss the benefits and costs of purchasing life insurance on the primary earners in a household.
West Virginia	12-3a. Explain why homeownersâ€™ insurance is required by a lender when a homeowner takes out a mortgage.
Wisconsin	12-10d. Compare the costs and benefits associated with for-profit versus non-profit credit counseling services.

Appendix C - *Median Cosine Similarity Score for Each Learning Objective*

Managing risk	
Standard	**Median Score**
12-1a. Discuss whether a premium paid to insure against a crash that never happens is wasted.	**0.58**
12-1b. Analyze the conditions under which it is appropriate for young adults to have life, health, and disability insurance.	**0.72**
12-2a. Identify individual characteristics that influence insurance purchase decisions.	**0.745**
12-2b. Recommend types of insurance needed by people with different characteristics.	**0.24**
12-3a. Explain why homeownersâ€™ insurance is required by a lender when a homeowner takes out a mortgage.	**0.255**
12-3b. Discuss why most states mandate auto liability coverage.	**0.495**
12-3c. Research the minimum auto liability insurance required in the state they live in and whether it is sufficient to cover typical auto accident financial losses.	**0.4**
12-4a. Research factors that result in lower auto insurance premiums.	**0.195**
12-4b. Explain why taking a safe driving course can lower a driverâ€™s auto insurance premium.	**0.24**
12-4c. Discuss the pros and cons of buying an auto insurance policy with a higher deductible.	**0.535**
12-5a. Discuss the advantages of obtaining health insurance coverage through an employer plan versus buying private insurance or being uninsured.	**0.92**
12-5b. Compare the cost of health insurance to the potential financial consequences of not having health insurance.	**0.445**
12-5c. Estimate the effect on different health insurance deductibles and coinsurance rates on out-of-pocket medical costs.	**0.39**
12-6a. Compare disability coverage offered by individual policies, employee benefit plans, Social Security, workersâ€™ compensation, and temporary disability programs (in some states).	**0.57**
12-6b. Assess the extent of financial risk and need for disability insurance using hypothetical disability scenarios.	**0.365**
12-7a. Explain the primary types of losses covered by auto, homeownerâ€™s, and renterâ€™s insurance policies.	**0.33**
12-7b. Describe situations where someone may be liable for injuries or damages to another person or their property.	**0.67**
12-7c. Identify factors that influence the cost of renterâ€™s insurance and homeownersâ€™ insurance.	**0.62**
12-8a. Explain how a personâ€™s death can result in financial losses to others.	**0.545**
12-8b. Discuss the benefits and costs of purchasing life insurance on the primary earners in a household.	**0.18**
12-9a. Discuss how state unemployment programs can help reduce economic hardship caused by job losses during a recession or pandemic.	**0.195**
12-9b. Compare the Medicare and Medicaid programs based on who they cover and how they are funded.	**0.435**
12-10a. Provide examples of insurance fraud.	**0.395**
12-10b. Investigate the legal consequence for individuals who are convicted of insurance fraud.	**0.73**
12-11a. Provide examples of how online behavior, e-mail and text-message scams, telemarketers, and other methods make consumers vulnerable to privacy infringement, identity theft, and fraud.	**0.475**
12-11b. Describe conditions under which individuals should and should not disclose their Social Security numbers, account numbers, or other sensitive information.	**0.85**
12-11c. Recommend strategies to reduce the risk of identity theft and financial fraud.	**0.63**
12-11d. Explain the steps an identity theft victim should take to limit losses and restore personal security.	**0.405**
12-12a. Evaluate the costs and benefits of buying an extended warranty on a specific item (e.g. cellphone, laptop, or vehicle) considering the likelihood of product failure, cost of replacing the item, and price of the warranty.	**0.24**
12-12b. Explain how extended warranties or service contracts are similar to and different from insurance.	**0.58**

Earning Income	
Standard	**Median Score**
12-1a. Research potential income and employee benefit packages that are likely to be offered to new employees by various companies, government agencies, or not-for-profit organizations.	**0.435**
12-1b. Explain why people should evaluate employee benefits in addition to wages and salaries when choosing between job and career opportunities.	**0.595**
12-1c. Differentiate between contributory and non-contributory employee benefits.	**0.5**
12-1d. Examine the benefits of participating in employer-sponsored retirement savings plans and healthcare savings plans.	**0.64**
12-2a. Give examples of intangible job benefits.	**0.375**
12-2b. Describe how intangible benefits can affect a workerâ€™s career choices and income.	**0.72**
12-2c. Evaluate the tradeoffs between income and non-income factors when making career or job choices.	**0.745**
12-3a. Evaluate the costs and benefits of investing in additional education or training.	**0.75**
12-3b. Explain how differences in peopleâ€™s life circumstances can affect their opportunity and willingness to further their education or training.	**0.725**
12-3c. Compare earnings and unemployment rates by level of education and training.	**0.605**
12-4a. Identify different types of jobs and careers where wages and salaries depend on a workerâ€™s productivity and skills.	**0.46**
12-4b. Explain why wages or salaries vary among employees in different types of jobs and among workers in the same jobs.	**0.38**
12-4c. Discuss possible explanations for the persistence of race and gender pay gaps.	**0.09**
12-5a. Discuss how economic and labor market conditions can affect income, career opportunities, and employment status.	**0.665**
12-5b. Evaluate the impact of technological advances on employment and income.	**0.61**
12-5c. Discuss the effects of an economic downturn on employment opportunities for people with different characteristics, such as education, experience, employment type, ethnicity, and gender.	**0.48**
12-6a. Calculate the amount of taxes a person is likely to pay when given information or data about the personâ€™s sources of income and amount of spending.	**0.56**
12-6b. Identify which level(s) of government typically receive(s) the tax revenue for income taxes, payroll taxes, property taxes, and sales taxes.	**0.44**
12-6c. Describe the benefits they receive, or may receive in the future, from government-collected tax revenue.	**0.38**
12-7a. Investigate the federal and state tax rates applicable to different sources of income.	**0.51**
12-7b. Compare sales tax rates paid on different types of goods in their state and for online purchases.	**0.33**
12-7c. Differentiate between gross, net, and taxable income.	**0.625**
12-7d. Explain why some income is reported on an IRS Form W-2 and some is reported on an IRS Form 1099, and how that could affect their taxes.	**0.465**
12-8a. Explain the difference between earned and unearned income.	**0.485**
12-8b. Compare the tax rates assessed on earned income, interest income, and capital gains income.	**0.625**
12-9a. Complete IRS Form W-4.	**0.26**
12-9b. Explain the difference between a tax credit and a tax deduction.	**0.405**
12-9c. Identify several examples of tax credits, determining whether they are refundable or non-refundable, and the groups of people who benefit most from each type.	**0.4**
12-10a. Identify different potential sources of retirement income.	**0.69**
12-10b. Describe the importance of having multiple sources of income in retirement, such as Social Security, employer-sponsored retirement plans, and personal investments.	**0.695**
12-10c. Explain the importance of participating in employer-sponsored retirement plans, when available, and contributing enough to qualify for the maximum employer match.	**0.41**
12-10d. Report the average benefit paid to a retiree living on Social Security today.	**0.25**
12-11a. Evaluate the benefits and costs of gig employment, such as driving for a cab or delivery service.	**0.305**
12-11b. Discuss the pros and cons of small business ownership as their primary source of income.	**0.34**

Spending	
Standard	**Median Score**
12-1a. Identify their short-term and long-term financial goals.	**1**
12-1b. Develop a budget to allocate current income to necessary and desired spending, including estimates for both fixed and variable expenses.	**0.64**
12-1c. Explain methods for adjusting a budget for unexpected expenses or emergencies.	**0.48**
12-1d. Evaluate the advantages of using budgeting tools, such as spreadsheets or apps.	**0.505**
12-2a. Select a product or service and describe the various factors that may influence a consumerâ€™s purchase decision.	**0.59**
12-2b. Describe a process for making an informed consumer decision.	**0.66**
12-2c. List the positive and negative effects of a recent consumer decision on the environment, society, and the economy.	**0.5**
12-3a. Explain the factors to evaluate when buying a durable good.	**0.49**
12-3b. Analyze the cost and features of three competing products or services.	**0.45**
12-3c. Compare product choices based on their impacts on the environment or society.	**0.535**
12-4a. List different ways retailers advertise the prices of their products.	**0.325**
12-4b. Describe how inflation affects purchase decisions and the price of goods and services.	**0.43**
12-4c. Summarize how negotiation affects consumer decisions and the price of goods and services.	**0.465**
12-5a. Explain how pre-purchase research encourages consumers to avoid impulse buying.	**0.56**
12-5b. Brainstorm consumer research strategies and resources to use when making purchase decisions.	**0.675**
12-5c. Analyze social media marketing and advertising techniques designed to encourage spending.	**0.415**
12-6a. Identify financial and personal reasons that younger adults often choose to rent a home instead of buying.	**0.505**
12-6b. Compare the short-term and long-term costs and benefits of renting versus buying a home in their city of residence.	**0.67**
12-6c. Define key rental contract terminology, including lease term, security deposit, grace period, and eviction.	**0**
12-7a. Discuss the motivations for and benefits of donating money, items, or time.	**0.445**
12-7b. Develop a list of charitable organizations and provide a possible reason that a donor might want to give money to each organization.	**0.265**
12-7c. Identify specific steps one should take when researching charitable and other not-for-profit organizations.	**0.305**
12-8a. Describe the roles and responsibilities of government agencies that help protect consumers from fraud.	**0.545**
12-8b. Identify state and federal consumer protection laws based on the issues they address and the safeguards they provide.	**0.59**
12-8c. Investigate common types of consumer fraud and unfair or deceptive business practices, including online scams, phone solicitations, and redlining.	**0.52**
12-8d. Make recommendations for sources of help for consumers who have experienced fraud.	**0.6**
12-9a. Explain how having a system for financial record-keeping can make it easier to make financial decisions.	**0.48**
12-9b. Develop a system for keeping track of spending, saving, and investing.	**0.73**
12-9c. Research financial technology options for financial record-keeping.	**0.46**

Savings	
Standard	**Median Score**
12-1a. Compare the features of regular savings accounts, money market accounts, and CDs.	**0.29**
12-1b. Explain why CDs typically pay higher interest rates than regular savings accounts or interest-bearing checking accounts.	**0.65**
12-2a. Select a preferred location for a savings account based on comparison of interest rates and fees at different types of financial institutions.	**0.16**
12-2b. Explain why an increase in the number of people who want to borrow money might result in banks paying higher rates on deposits.	**0.33**
12-2c. Discuss types of market conditions that could result in financial institutions paying lower rates on savings accounts.	**0.36**
12-3a. Research mobile payment account alternatives.	**0.59**
12-3b. Compare and contrast the features of mobile payment accounts, cryptocurrency accounts, and checking/savings accounts.	**0.365**
12-3c. Explain why storing money in a mobile payment account can reduce the ability to grow savings.	**0.195**
12-4a. Explain why savers typically earn a higher nominal rate of interest when inflation is high.	**0.285**
12-4b. Illustrate how inflation can reduce the purchasing power of savings over time if the nominal interest rate is lower than the inflation rate.	**0.375**
12-4c. Investigate how federal I bonds provide inflation protection for savers.	**0.435**
12-5a. Investigate the areas of financial institution operations that are subject to state and/or federal regulation and supervision.	**0.48**
12-5b. Identify the state agency responsible for regulating financial institutions where they live.	**0.26**
12-5c. Explain the importance of solvency regulation for financial institutions.	**0.41**
12-6a. Explain how traditional IRAs (individual retirement accounts), Roth IRAs, and education savings accounts provide incentives for people to save.	**0.505**
12-6b. Compare the tax advantages of traditional and Roth IRAs.	**0.49**
12-6c. Compare the tax advantages of different types of education savings accounts.	**0.525**
12-7a. Explain how an employer match of employee contributions to its retirement plan provides an incentive for employees to save.	**0.48**
12-7b. Compare the impact of employee â€œopt inâ€ versus â€œopt outâ€ of employer retirement plans and explain why it makes a difference.	**0.29**
12-7c. Describe the pros and cons of saving through an employer retirement plan as compared to saving outside of an employer plan.	**0.38**
12-7d. Explain the benefits of saving money in a health savings account for individuals with high-deductible health plans.	**0.73**
12-8a. Assess the value of sharing financial goals and personal financial information with a partner before combining finances.	**0.675**
12-8b. Discuss how personal financial decisions can affect other people.	**0.56**
12-9a. Explain how external influences (e.g. peers, family, or social media) can impact personal savings decisions.	**0.36**
12-9b. Identify strategies to manage psychological and emotional obstacles to saving.	**0.595**
12-9c. Discuss strategies for avoiding personal triggers that result in deviating from a savings plan.	**0.665**
12-9d. Explain how the saving strategy â€œpay yourself firstâ€ can help people achieve their saving goals.	**0.98**

Investing	
Standard	**Median Score**
12-1a. Give examples of factors that can influence a person's risk tolerance.	**0.64**
12-1b. Discuss how a person's risk tolerance influences their investment decisions.	**0.365**
12-1c. Assess their personal risk tolerance using an online tool or worksheet.	**0.27**
12-2a. Describe the different types of annual cash flows that can be received by investors.	**0.445**
12-2b. Compare nominal annual rates of return over time on different types of investments, including cash flows and price changes.	**0.355**
12-2c. Explain why assets that do not produce income or are exposed to large price fluctuation (such as collectibles, precious metals, and cryptocurrencies) are described as speculative investments.	**0.665**
12-3a. Discuss the advantages and disadvantages of investing in riskier assets.	**0.56**
12-3b. Investigate the long-run average rates of returns on small-company stocks, large-company stocks, corporate bonds, and Treasury bonds.	**0.405**
12-3c. Explain why the expected rate of return on a value stock or mutual fund is likely to be lower than that of a growth stock or mutual fund.	**0.31**
12-3d. Explain why bonds with longer maturities generally earn a higher return than shorter-term bonds.	**0.325**
12-4a. Describe the impact of inflation on prices over time.	**0.395**
12-4b. Explain the relationship between nominal and real returns.	**0.31**
12-4c. Find the current rate paid on CDs at a bank and calculate the expected real rate after inflation.	**0.67**
12-5a. Describe factors that influence the prices of financial assets.	**0.25**
12-5b. Predict what could happen to the price of a stock if new information is reported about the company or its products.	**0.45**
12-5c. Discuss how economic downturns that result in high unemployment can affect the prices of financial assets.	**0.315**
12-5d. Explain why the market price of some assets, such as bonds and real estate, increase when interest rates decrease.	**0.72**
12-6a. Recommend portfolio allocation between major asset classes for a short-term goal versus a long-term goal.	**0.605**
12-6b. Discuss the pros and cons of investing in a diversified mutual fund versus investing in a small number of individual stocks.	**0.435**
12-6c. Suggest an appropriate asset allocation for a very risk averse person versus a very risk tolerant person.	**0.45**
12-6d. Explain how target date retirement funds reallocate investments over time to meet their investment objective.	**0.64**
12-7a. Discuss how the expenses associated with buying and selling investments can impact rates of return and investment outcomes.	**0.47**
12-7b. Compare the expense ratios for several mutual funds.	**0.16**
12-7c. Explain why an actively managed mutual fund usually has a higher expense ratio than an index fund.	**0.51**
12-8a. Compare tax rates paid on interest income versus short-term and long-term capital gains.	**0.535**
12-8b. Describe the advantages of investing through a tax-deferred account such as an IRA or 401(k) versus a taxable account.	**0.42**
12-8c. Investigate the contribution limits and tax advantages of a traditional IRA versus a Roth IRA.	**0.45**
12-9a. Identify several behavioral biases that can result in poor investment decisions (e.g. loss aversion, investing in employer stock, home bias, mental accounting).	**0.23**
12-9b. Brainstorm methods for avoiding negative consequences from behavioral biases.	**0.41**
12-10a. Explore common financial technologies used for investing, including automated trading platforms.	**0.415**
12-10b. Explain how automating investment activities can help people avoid making emotional investment decisions.	**0.49**
12-11a. Discuss how the development of financial technology has made it easier for people of all income and education levels to participate in financial markets.	**0.285**
12-11b. Choose a discount broker and research the minimum starting account balance, minimum monthly investment, and trading costs.	**0.47**
12-11c. Identify the advantages and disadvantages of robo-advising and other investment-related financial technologies.	**0.365**
12-12a. Explain the role of federal regulators in financial markets.	**0.34**
12-12b. Discuss why insider trading is illegal and harmful to investment markets.	**0.535**
12-12c. Explain the importance of having access to full and accurate information about potential investments.	**0.355**
12-13a. Explain why investors often compare portfolio performance to a benchmark such as the S&P 500 Index.	**0.165**
12-13b. Research the composition of the most popular benchmark indices and compare their recent performance.	**0.505**
12-13c. Discuss the advantages of investing in an exchange-traded fund (ETF) that tracks a market index rather than investing in actively managed mutual funds or individual stocks and bonds.	**0.75**
12-14a. Discuss reasons that a person might want to hire a financial professional to manage their investments or provide investment advice.	**0.41**
12-14b. Explain the importance of licensing, certifications, education, and experience as criteria for selecting a financial professional for investment management or advice.	**0.495**
12-14c. Investigate where and how to find qualified financial professionals.	**0.25**

Managing Credit	
Standard	**Median Score**
12-1a. Describe how credit card grace periods, methods of interest calculation, and fees affect borrowing costs.	**0.64**
12-1b. Compare the cost of borrowing $1,000 using consumer credit options that differ in rates and fees.	**0.355**
12-2a. Give examples of unsecured and secured loans.	**0.335**
12-2b. Explain why lenders charge lower interest rates on secured loans than on unsecured loans.	**0.605**
12-2c. Compare what happens if a borrower fails to make required payments on a secured loan, such as an auto loan or a home mortgage, versus failing to pay a credit card account.	**0.345**
12-3a. Identify the type of collateral required for a mortgage loan.	**0.47**
12-3b. Differentiate between adjustable-rate and fixed-rate mortgages.	**0.56**
12-3c. Compare monthly mortgage payments for loans that differ in repayment period, amount borrowed, and interest rate.	**0.505**
12-4a. Describe the different sources of funding for post-secondary education.	**0.245**
12-4b. Explain the role the FAFSA plays in applying for college financial aid.	**0.27**
12-4c. Identify scholarships and grants for which they are eligible.	**0.34**
12-4d. Estimate the reduction in total cost of education and potential student loan debt if they complete their first two years of college at a community college before transferring to a four-year institution.	**0.395**
12-5a. Compare federal and private student loans based on interest rates, repayment rules, and other characteristics.	**0.28**
12-5b. Describe the process of applying for a student loan.	**0.38**
12-5c. Estimate total interest on various student loans based on interest rates and repayment plans.	**0.41**
12-5d. Predict the potential consequences of deferred payment of student loans.	**0.43**
12-6a. Identify examples of loans that may require down payments.	**0.215**
12-6b. Given the price of a home, estimate the amount of down payment required.	**0.27**
12-6c. For a specified loan amount, compare the monthly loan payment with a 10% down payment versus a 20% down payment.	**0.395**
12-6d. Explain how a down payment makes a borrower more attractive to a lender and motivates loan repayment by the borrower.	**0.495**
12-7a. Identify the primary organizations that maintain and provide consumer credit reports.	**0.52**
12-7b. Assess the value to a potential lender of the information contained in a credit report.	**0.55**
12-7c. Explain how a person can get a free copy of their credit report and why this is advisable.	**0.385**
12-7d. Outline the process of disputing inaccurate credit report information.	**0.525**
12-8a. Identify the main factors that are included in credit score calculations.	**0.785**
12-8b. Explain how a borrowerâ€™s credit score can impact their cost of credit and their ability to get credit.	**0.55**
12-8c. Recommend ways that a person can increase their credit score.	**0.525**
12-9a. Explain how landlords, potential employers, and insurance companies use credit reports and credit scores in decision-making.	**0.655**
12-9b. Provide examples of benefits associated with having a good credit score.	**0.51**
12-9c. Compare the effect of soft versus hard credit inquiries on a personâ€™s credit score.	**0.505**
12-10a. Describe how failing to repay a loan can negatively impact a personâ€™s finances and life.	**0.52**
12-10b. Identify sources of assistance with debt management.	**0.575**
12-10c. Create a plan for a person who is having difficulty repaying debt.	**0.485**
12-10d. Compare the costs and benefits associated with for-profit versus non-profit credit counseling services.	**0.195**
12-11a. Describe the purpose of bankruptcy laws.	**0.595**
12-11b. Investigate the effects of bankruptcy on assets, employment, and future access to credit.	**0.14**
12-11c. Compare the results of liquidation versus reorganization bankruptcy.	**0.38**
12-12a. Explain the rationale behind laws that require people to have access to full information about credit cards and loans before they borrow money.	**0.27**
12-12b. Discuss the importance of protecting borrowers from discrimination and abusive marketing or collection practices.	**0.355**
12-12c. Research where to find credible sources of up-to-date information on credit rights and responsibilities.	**0.375**
12-13a. Identify products and practices that are classified as alternative financial services.	**0.515**
12-13b. Discuss the costs and benefits of using alternative financial services relative to traditional banking.	**0.28**
12-13c. Explain how using payday loans can cause a cycle of debt.	**0.41**

Appendix D - *Simple Random Sample for Confidence Interval*

To show all 5000+ learning objective cosine similarity scores would require hundreds of pages in the document, so only the simple random sample of 100 learning objectives is provided.

1	State	Category	Cosine Similarity	National Standard	State Standard	Human Score
2	tn	credit	0.48	12-9c. Compare the effect of soft versus hard credit inquiries on a personâ€™s credit score.	Describe credit reports and credit scores. Describe the relationship between consumers and credit reports/credit scores. Explain how the credit score may impact borrowing opportunities and the cost of credit. Summarize specific activities used to maintain a good credit score.	1
3	ne	credit	0.15	12-12b. Discuss the importance of protecting borrowers from discrimination and abusive marketing or collection practices.	SS 12.2.8.c Compare and contrast the cost and benefits of various lending institutions	1
4	ga	credit	0.48	12-3b. Differentiate between adjustable-rate and fixed-rate mortgages.	Define annual percentage rate and analyze how different interest rates can affect monthly payments on loans.	1
5	ks	managing risk	1	12-7c. Identify factors that influence the cost of renterâ€™s insurance and homeownersâ€™ insurance.	12-7c. Identify factors that influence the cost of renterâ€™s insurance and homeownersâ€™ insurance.	1
6	mn	investing	0.44	12-4b. Explain the relationship between nominal and real returns.	9.2.9.3 Evaluate how various household assets of property, housing, stocks, bonds, savings accounts and monetizing creative enterprises can generate income, considering risk, return and time horizon.	1
7	ri	investing	0.31	12-11a. Discuss how the development of financial technology has made it easier for people of all income and education levels to participate in financial markets.	8-3c. Discuss how specific skills training can improve a young personâ€™s human capital, productivity, and income-earning potential.	1
8	nc	saving	0.21	12-5b. Identify the state agency responsible for regulating financial institutions where they live.	EPF.MCM.2.1 Design a plan that uses the services of various financial institutions to meet financial goals.	1
9	mo	credit	0.54	12-7c. Explain how a person can get a free copy of their credit report and why this is advisable.	E. Evaluate a credit report to verify accuracy. (9-12)	1
10	mi	credit	0.05	12-11c. Compare the results of liquidation versus reorganization bankruptcy.	Financial Investing â€“ analyze the risks, expected rate of return, tax benefits, impact of inflation, role of government agencies, and importance of diversification when investing in financial assets.	0
11	oh	spending	0.09	12-6c. Define key rental contract terminology, including lease term, security deposit, grace period, and eviction.	19. Credit is a contractual agreement in which a borrower receives something of value now and agrees to repay to lender at some later date.	0
12	fl	managing risk	0.9	12-9a. Discuss how state unemployment programs can help reduce economic hardship caused by job losses during a recession or pandemic.	SS.912.FL.1.5 Discuss reasons why changes in economic conditions or the labor market can cause changes in a workerâ€™s income or may cause unemployment.	1
13	ia	credit	0.35	12-9b. Provide examples of benefits associated with having a good credit score.	SS-FL.9-12.18. Analyze the cost and benefits of different types of credit and debt. (9-12)	1
14	pa	managing risk	0.32	12-7b. Describe situations where someone may be liable for injuries or damages to another person or their property.	Explain how consumer rights and responsibilities are protected.	1
15	ca	earning income	1	12-3b. Explain how differences in peopleâ€™s life circumstances can affect their opportunity and willingness to further their education or training.	12-3b. Explain how differences in peopleâ€™s life circumstances can affect their opportunity and willingness to further their education or training.	1
16	or	saving	0.49	12-9c. Discuss strategies for avoiding personal triggers that result in deviating from a savings plan.	2.17. Explain various methods of saving and how saving can help reach financial goals. (2)	1
17	nh	saving	0.07	12-7d. Explain the benefits of saving money in a health savings account for individuals with high-deductible health plans.	SS:EC:8:6.3: Demonstrate the use of the different types of accounts available from financial institutions, e.g., checking or savings accounts. (7, 8)	0
18	pa	investing	0.5	12-2b. Compare nominal annual rates of return over time on different types of investments, including cash flows and price changes.	Analyze the management of financial resources across the lifespan.	1
19	wi	investing	1	12-12a. Explain the role of federal regulators in financial markets.	SI2.e.i Explain how federal and state regulators help protect investors. Identify options that are tax free. (3-5)	1

20	ms	earning income	0.11	12-6a. Calculate the amount of taxes a person is likely to pay when given information or data about the personâ€™s sources of income and amount of spending.	1. Compare sources of personal income and compensation and analyze factors that affect net income.	0
21	ia	saving	0.15	12-5b. Identify the state agency responsible for regulating financial institutions where they live.	SS-FL.9-12.19. Summarize a borrower's rights and responsibilities. (9-12)	1
22	la	managing risk	0.12	12-3b. Discuss why most states mandate auto liability coverage.	c. Explain the benefits and risks of using credit and examine the various uses.	1
23	ca	credit	1	12-4b. Explain the role the FAFSA plays in applying for college financial aid.	12-4b. Explain the role the FAFSA plays in applying for college financial aid.	1
24	ut	spending	0.56	12-1b. Develop a budget to allocate current income to necessary and desired spending, including estimates for both fixed and variable expenses.	4.1.1 Develop a budget.	1
25	mi	managing risk	0.74	12-8a. Explain how a personâ€™s death can result in financial losses to others.	Protecting and Insuring â€“ assess the financial risk of lost income, assets, health, or identity, and determine if a person should accept the risk exposure, reduce risk, or transfer the risk to others by paying a fee now to avoid the possibility of a larger loss later.	1
26	ga	investing	1	12-8b. Describe the advantages of investing through a tax-deferred account such as an IRA or 401(k) versus a taxable account.	Explain the advantages of using tax-advantaged retirement planning including a tradition Individual Retirement Account (IRA), a Roth IRA, a myRA, and a company 401K or 403b.	1
27	sc	credit	0.08	12-6c. For a specified loan amount, compare the monthly loan payment with a 10% down payment versus a 20% down payment.	Complete a loan application.	0
28	la	credit	0.61	12-10a. Describe how failing to repay a loan can negatively impact a personâ€™s finances and life.	e. Create a budget and explain its importance in achieving personal financial goals and avoiding negative financial consequences.	1
29	al	saving	0.45	12-7b. Compare the impact of employee â€œopt inâ€ versus â€œopt outâ€ of employer retirement plans and explain why it makes a difference.	d. Compare and contrast retirement and estate planning options available.	1
30	mi	saving	0.11	12-3b. Compare and contrast the features of mobile payment accounts, cryptocurrency accounts, and checking/savings accounts.	Saving â€“ identify the incentives people have to set aside income for future consumption, and evaluate the impact of time, interest rates, and inflation upon the value of savings.	0
31	ks	earning income	1	12-3a. Evaluate the costs and benefits of investing in additional education or training.	12-3a. Evaluate the costs and benefits of investing in additional education or training.	1
32	mn	credit	0.07	12-7d. Outline the process of disputing inaccurate credit report information.	9.2.9.4 Evaluate the benefits and costs of credit. Explain how the financial industry assesses an individual's ability to manage credit and how this affects their ability to borrow, rent, get a job, and achieve other financial goals.	0
33	mn	investing	0.4	12-10a. Explore common financial technologies used for investing, including automated trading platforms.	9.2.9.3 Evaluate how various household assets of property, housing, stocks, bonds, savings accounts and monetizing creative enterprises can generate income, considering risk, return and time horizon.	1
34	ri	spending	1	12-7c. Identify specific steps one should take when researching charitable and other not-for-profit organizations.	12-7c. Identify specific steps one should take when researching charitable and other not-for-profit organizations.	1
35	pa	investing	0.62	12-7a. Discuss how the expenses associated with buying and selling investments can impact rates of return and investment outcomes.	Compare the influences of income and fringe benefits to make decisions about work.	1
36	al	credit	0.21	12-10d. Compare the costs and benefits associated with for-profit versus non-profit credit counseling services.	a. Examine and report on the types of services offered by consumer credit counseling agencies.	1
37	fl	managing risk	0	12-9b. Compare the Medicare and Medicaid programs based on who they cover and how they are funded.	SS.912.FL.6.7 Compare the purposes of various types of insurance.	0
38	ia	managing risk	0.6	12-2b. Recommend types of insurance needed by people with different characteristics.	SS.8.29. Identify ways insurance may minimize personal financial risk. (8)	1
39	fl	credit	0.77	12-2c. Compare what happens if a borrower fails to make required payments on a secured loan, such as an auto loan or a home mortgage, versus failing to pay a credit card account.	SS.912.FL.4.8 Examine the fact that failure to repay a loan has significant consequences for borrowers such as negative entries on their credit report, repossession of property (collateral), garnishment of wages, etc.	1
40	wi	investing	0.66	12-13c. Discuss the advantages of investing in an exchange-traded fund (ETF) that tracks a market index rather than investing in actively managed mutual funds or individual stocks and bonds.	SI2.b.m Explore investing choices (e.g., collectibles, stocks, bonds, or mutual funds) which can produce income or growth. Identify the differences between banks, credit unions, and investment firms. (6-8)	1

41	mo	credit	0.09	12-11c. Compare the results of liquidation versus reorganization bankruptcy.	B. Compare the liquidity, interest payment or penalty of various savings instruments. (9-12)	0
42	ia	saving	0.28	12-4b. Illustrate how inflation can reduce the purchasing power of savings over time if the nominal interest rate is lower than the inflation rate.	S.6.25. Demonstrate how to allocate income for spending, saving and giving. (6)	1
43	oh	credit	0.41	12-8a. Identify the main factors that are included in credit score calculations.	23. Credit and debt affect tax obligations.	1
44	ri	earning income	0.09	12-4c. Discuss possible explanations for the persistence of race and gender pay gaps.	12-5a. Discuss how economic and labor market conditions can affect income, career opportunities, and employment status.	0
45	al	credit	0.27	12-7d. Outline the process of disputing inaccurate credit report information.	a. Investigate and report on sources of assistance for resolving consumer disputes.	1
46	oh	investing	0.45	12-5a. Describe factors that influence the prices of financial assets.	17. Investment strategies must take several factors into consideration.	1
47	ne	earning income	0.63	12-6c. Describe the benefits they receive, or may receive in the future, from government-collected tax revenue.	SS 12.2.7.c Assess the effects of taxes on personal income	1
48	or	earning income	0.42	12-1c. Differentiate between contributory and non-contributory employee benefits.	1.19. Identify sources of income (e.g., gifts, borrowing, allowance, work wages). (1)	1
49	ia	spending	0.49	12-6a. Identify financial and personal reasons that younger adults often choose to rent a home instead of buying.	SS.6.24. Explain how personal financial decisions are influenced by an individualâ€™s needs and wants. (6)	1
50	tn	credit	0.32	12-3a. Identify the type of collateral required for a mortgage loan.	Compare and contrast various types of credit and calculate the real cost of borrowing. Explain factors that can affect the approval process associated with each type. Identify information and procedures required in the credit application process. Analyze factors associated with the purchase of an automobile and defend a specific buying decision.	1
51	ct	managing risk	0	12-12b. Explain how extended warranties or service contracts are similar to and different from insurance.	4-4b. Investigate the types of insurance commonly available for people to purchase.	0
52	va	investing	0.31	12-4a. Describe the impact of inflation on prices over time.	examining the effect of supply and demand on wages and prices.	1
53	ct	earning income	0.05	12-11a. Evaluate the benefits and costs of gig employment, such as driving for a cab or delivery service.	8-9a. Investigate the motivating factors to being self-employed or working as an independent contractor in the â€œgigâ€ economy.	0
54	va	spending	0.12	12-5c. Analyze social media marketing and advertising techniques designed to encourage spending.	evaluating discretionary spending decisions.	1
55	la	earning income	0.9	12-3a. Evaluate the costs and benefits of investing in additional education or training.	a. Explain the relationship between education, training, and career options to future earning potential.	1
56	ct	saving	0.02	12-2b. Explain why an increase in the number of people who want to borrow money might result in banks paying higher rates on deposits.	4-3a. Explain why a person might prefer to lend an item or money to one person over another.	0
57	nc	earning income	0.49	12-1a. Research potential income and employee benefit packages that are likely to be offered to new employees by various companies, government agencies, or not-for-profit organizations.	EPF.IE.1.3 Identify the costs of postsecondary education and the potential increase in income from a career of choice.	1
58	in	earning income	0	12-9a. Complete IRS Form W-4.	b. Calculate the amount of taxes on investments and income tax-free earnings	0
59	la	investing	0.31	12-5c. Discuss how economic downturns that result in high unemployment can affect the prices of financial assets.	e. Create a budget and explain its importance in achieving personal financial goals and avoiding negative financial consequences.	1
60	ga	managing risk	0.52	12-11a. Provide examples of how online behavior, e-mail and text-message scams, telemarketers, and other methods make consumers vulnerable to privacy infringement, identity theft, and fraud.	Describe common ways identity theft happens.	1
61	ne	saving	0.14	12-1b. Explain why CDs typically pay higher interest rates than regular savings accounts or interest-bearing checking accounts.	SS 12.2.8.c Compare and contrast the cost and benefits of various lending institutions	1
62	ga	saving	0.22	12-1b. Explain why CDs typically pay higher interest rates than regular savings accounts or interest-bearing checking accounts.	Compare interest rates on loans and credit from institutions.	1
63	wv	investing	0.39	12-10b. Explain how automating investment activities can help people avoid making emotional investment decisions.	Investigate the consequences of personal choices in relation to finances.	1

64	oh	credit	0.47	12-10d. Compare the costs and benefits associated with for-profit versus non-profit credit counseling services.	13. Utilizing financial services and risk management tools, and interpreting and comparing consumer lending statements, terms and conditions enable one to be an informed consumer.	1
65	ms	earning income	0.34	12-4a. Identify different types of jobs and careers where wages and salaries depend on a workerâ€™s productivity and skills.	c. Compare the unemployment rates of workers with different levels of education.	1
66	ca	investing	0.23	12-10a. Explore common financial technologies used for investing, including automated trading platforms.	12-9c. Research financial technology options for financial record-keeping.	1
67	in	credit	0.29	12-13b. Discuss the costs and benefits of using alternative financial services relative to traditional banking.	a. Compare advantages and disadvantages of buying and selling investments through various channels, including financial advisors, investment clubs, and online brokers	1
68	ri	managing risk	1	12-11b. Describe conditions under which individuals should and should not disclose their Social Security numbers, account numbers, or other sensitive information.	12-11b. Describe conditions under which individuals should and should not disclose their Social Security numbers, account numbers, or other sensitive information.	1
69	ms	managing risk	0.14	12-9b. Compare the Medicare and Medicaid programs based on who they cover and how they are funded.	e. Explain the fundamentals of health insurance and identify various types of health insurance coverage.	1
70	ks	earning income	0.12	12-7a. Investigate the federal and state tax rates applicable to different sources of income.	12-7c. Differentiate between gross, net, and taxable income.	1
71	ca	investing	0.2	12-11c. Identify the advantages and disadvantages of robo-advising and other investment-related financial technologies.	8-1a. List the potential benefits of investing money in a financial asset.	1
72	nc	credit	0.23	12-2b. Explain why lenders charge lower interest rates on secured loans than on unsecured loans.	EPF.MCM.2.2 Explain how interest and fees impact spending, debt, and savings.	1
73	ga	investing	0.72	12-6a. Recommend portfolio allocation between major asset classes for a short-term goal versus a long-term goal.	Describe the differences in strategies used for long-term investing vs. short term investing.	1
74	ga	investing	0.29	12-11a. Discuss how the development of financial technology has made it easier for people of all income and education levels to participate in financial markets.	Explain that some financial institutions are for profit and others are non-profit.	1
75	ms	investing	0.19	12-5d. Explain why the market price of some assets, such as bonds and real estate, increase when interest rates decrease.	b. Explain the effect of debt on net worth and the ability to borrow money.	1
76	ks	credit	1	12-3b. Differentiate between adjustable-rate and fixed-rate mortgages.	12-3b. Differentiate between adjustable-rate and fixed-rate mortgages.	1
77	wi	spending	0.45	12-7c. Identify specific steps one should take when researching charitable and other not-for-profit organizations.	FM2.e.i Describe the benefits of charitable giving, and volunteerism. (3-5)	1
78	ct	managing risk	1	12-11d. Explain the steps an identity theft victim should take to limit losses and restore personal security.	12-11d. Explain the steps an identity theft victim should take to limit losses and restore personal security.	1
79	nc	saving	1	12-2a. Select a preferred location for a savings account based on comparison of interest rates and fees at different types of financial institutions.	EPF.MCM.2.3 Compare costs and benefits of cash, debit, payment applications, and credit card transactions in terms of interest rates, fees, penalties, costs, and benefits.	1
80	ut	saving	0.13	12-7a. Explain how an employer match of employee contributions to its retirement plan provides an incentive for employees to save.	2.2.4 Understand the use and advantages of 529 plans and the benefit of planning early for paying for the cost of post-secondary education and training.	1
81	ks	spending	0.18	12-7b. Develop a list of charitable organizations and provide a possible reason that a donor might want to give money to each organization.	12-7c. Identify specific steps one should take when researching charitable and other not-for-profit organizations.	1
82	ri	credit	1	12-8a. Identify the main factors that are included in credit score calculations.	12-8a. Identify the main factors that are included in credit score calculations.	1
83	wi	saving	0.2	12-2c. Discuss types of market conditions that could result in financial institutions paying lower rates on savings accounts.	SI1.b.i Describe why a person deposits money into a financial institution. Describe characteristics of a secure savings account. (3-5)	1
84	ks	saving	1	12-9c. Discuss strategies for avoiding personal triggers that result in deviating from a savings plan.	12-9c. Discuss strategies for avoiding personal triggers that result in deviating from a savings plan.	1
85	ne	managing risk	0.52	12-4a. Research factors that result in lower auto insurance premiums.	SS 12.2.8.a Analyze factors that affect credit	1
86	nh	saving	0.11	12-7a. Explain how an employer match of employee contributions to its retirement plan provides an incentive for employees to save.	SS:EC:8:6.5: Define and compare saving and investing. (7, 8)	0

86	nh	saving	0.11	12-7a. Explain how an employer match of employee contributions to its retirement plan provides an incentive for employees to save.	SS:EC:8:6.5: Define and compare saving and investing. (7, 8)	0
87	mi	saving	0.29	12-5a. Investigate the areas of financial institution operations that are subject to state and/or federal regulation and supervision.	Financial Investing â€“ analyze the risks, expected rate of return, tax benefits, impact of inflation, role of government agencies, and importance of diversification when investing in financial assets.	1
88	ia	earning income	0.5	12-1b. Explain why people should evaluate employee benefits in addition to wages and salaries when choosing between job and career opportunities.	SS.6.24. Explain how personal financial decisions are influenced by an individualâ€™s needs and wants. (6)	1
89	va	credit	0.07	12-3b. Differentiate between adjustable-rate and fixed-rate mortgages.	comparing the types of financial institutions.	0
90	nc	credit	0.36	12-6c. For a specified loan amount, compare the monthly loan payment with a 10% down payment versus a 20% down payment.	EPF.MCM.2.3 Compare costs and benefits of cash, debit, payment applications, and credit card transactions in terms of interest rates, fees, penalties, costs, and benefits.	1
91	ca	earning income	0.19	12-6c. Describe the benefits they receive, or may receive in the future, from government-collected tax revenue.	4-7a. Describe examples of government-provided goods and services that are paid for with taxes.	1
92	in	credit	0.39	12-7a. Identify the primary organizations that maintain and provide consumer credit reports.	a. Analyze online and printed resources for up-to-date information about consumer credit rights	1
93	mn	saving	0.34	12-6a. Explain how traditional IRAs (individual retirement accounts), Roth IRAs, and education savings accounts provide incentives for people to save.	9.2.9.2 Establish personal financial goals. Create a financial plan, considering budgeting and asset building to meet those goals. Determine the nature of the barriers the individual confronts. Determine ways to track the success of the plan.	1
94	mn	spending	0.27	12-1c. Explain methods for adjusting a budget for unexpected expenses or emergencies.	2.2.9.1 Given a goal and several alternative choices to reach that goal, select the best choice and explain.	1
95	mo	credit	0.82	12-9a. Explain how landlords, potential employers, and insurance companies use credit reports and credit scores in decision-making.	D. Analyze why credit scores may be used by entities such as employers and landlords. (9-12)	1
96	fl	investing	0.13	12-6d. Explain how target date retirement funds reallocate investments over time to meet their investment objective.	SS.4.FL.5.1 Explain that after people have saved some of their income, they must decide how to invest their savings.	1
97	mi	saving	0.53	12-7d. Explain the benefits of saving money in a health savings account for individuals with high-deductible health plans.	Protecting and Insuring â€“ assess the financial risk of lost income, assets, health, or identity, and determine if a person should accept the risk exposure, reduce risk, or transfer the risk to others by paying a fee now to avoid the possibility of a larger loss later.	1
98	mo	investing	0.36	12-2a. Describe the different types of annual cash flows that can be received by investors.	A. Explain how the rate of return earned from investments will vary according to the amount of risk. (9-12)	1
99	mn	managing risk	0.09	12-9a. Discuss how state unemployment programs can help reduce economic hardship caused by job losses during a recession or pandemic.	9.2.9.1 Analyze how individual or household income is determined by a variety of individual and social factors.	0
100	or	credit	0.81	12-2c. Compare what happens if a borrower fails to make required payments on a secured loan, such as an auto loan or a home mortgage, versus failing to pay a credit card account.	HS.43. Compare and contrast various types of loans available and how to obtain them, including student loans (9-12)	1
101	ct	earning income	1	12-2c. Evaluate the tradeoffs between income and non-income factors when making career or job choices.	12-2c. Evaluate the tradeoffs between income and non-income factors when making career or job choices.	1
102					**Total:**	83

Appendix E - Implementation and Technical Setup

This project was developed using over 150 lines of Python code across modular functions, each tailored for a specific role in an NLP pipeline. The overall process included extracting financial literacy standards from state and national curriculum PDFs, preprocessing the text for NLP readiness, and then embedding these standards into high-dimensional vectors using OpenAI's

text-embedding-ada-002 model. To facilitate this pipeline, the researcher leveraged open-source libraries, official documentation, and example code snippets from trusted developer forums (OpenAI, 2024; Liang, 2024). While some components were written from scratch, many implementations drew on widely used design patterns in the NLP community.

In the preprocessing phase, the fitz module from (Artifex, 2019) was used to extract tables from PDF curriculum files. This allowed accurate parsing of state standards into a structured tabular form. The data was then organized with pandas and cleaned using NLTK—a standard toolkit for natural language processing. Preprocessing steps included lowercasing, stopword removal, punctuation filtering, and lemmatization using WordNetLemmatizer. Custom domain-specific terms (e.g., "investment_bank", "credit_score") were preserved by incorporating a financial glossary (financial_terms.txt). Regular expressions (re) were employed to remove extraneous characters such as digits and special symbols, ensuring semantic clarity. These steps mirror best practices in NLP as documented in guides like (Mariya Mansurova, 2024).

Following the guide of Siladitya Ghosh in this process, embedding generation used OpenAI's Embedding API with the text-embedding-ada-002 model, which encodes text into 1536-dimensional semantic vectors (Ghosh, 2024). Embedding requests were formatted in JSON, allowing efficient parallel processing. Custom batch jobs were created using Python's json, os, textwrap, dedent, and string modules. For example, each entry was structured as: { "custom_id": "CEE-standard-01", "method": "POST", "url": "/v1/embeddings", "body": { "input": "text here", "model": "text-embedding-ada-002", "encoding_format": "float" } }. Cleaned data was also organized into dictionaries using (Collins, 2024), converting standards into a JSON structure suitable for further semantic alignment. Additionally, OpenAI's was used in extracting learning objectives, enforcing a strict JSON schema for response validation.

This pipeline demonstrates a modular, end-to-end approach to educational NLP, combining community knowledge with domain-specific logic. Code functions like extract_table(), df_to_dict(), clean_text(), preprocess_pdf(), and create_batch_file() were structured to allow flexible reuse and systematic transformation of raw curriculum data. The final architecture reflects current best practices in both AI-assisted research and reproducible data science workflows.

For transparency and replication, a complete repository of the code, including preprocessing routines, prompt schemas, glossary integration, and cosine similarity calculations, has been made publicly available on GitHub. Each function is documented with inline comments and example usage, and a sample .ipynb notebook is included to demonstrate the full workflow from PDF parsing to vector output and alignment scoring.

Appendix F - Ipynb Notebook

 # install necessary libraries

!pip install PyMuPDF

!pip install tools

 # create folders to save output

!mkdir pdfs

!mkdir json

```
!mkdir json_clean

!mkdir json_embeddings

#import necessary libraries

import fitz

import os

import pandas as pd

import string

import re

import json

from text wrap import dedent

from openai import OpenAI



# NLTK for preprocessing the data

import nltk

from nltk.tokenize import word_tokenize

from nltk.corpus import stopwords

from nltk.stem import WordNetLemmatizer
```

```
# Download required NLTK data
nltk.download('punkt')
nltk.download('punkt_tab')
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('averaged_perceptron_tagger')
# load financial terms from file
with open('financial_terms.txt', 'r') as f:
    lst_financial_terms = f.readlines()


FINANCIAL_TERMS = set([term.strip() for term in lst_financial_terms])
client = OpenAI(  api_key="removed for privacy reasons")


# function to extract table from the pdf
def extract_table(pdf_path):
    # open pdf using fitz
```

```
    doc = fitz.open(pdf_path)


    # create dataframe to save the table
    df = pd.DataFrame()


    # extract tables from multiple pages and combine into a single table
    for page in doc:
        tabs = page.find_tables()
        if tabs.tables:
            df = pd.concat([df, pd.DataFrame(tabs[0].extract())]) # combine the tables


    return df


# function to convert the table to the json format
def df_to_dict(df, is_national = False):
    if is_national:
```

```
    # convert the text in df column 0 to lower and remove all characters except a-z and remove all space characters with single space

    df = df[~(df[0].isna() | (df[0] == ''))]

    df.iloc[:, 0] = df[0].str.lower()

    df.iloc[:, 0] = df[0].apply(lambda x: re.sub(r'\s+', ' ', x))

    df.iloc[:, 0] = df[0].apply(lambda x: re.sub(r'[^a-z ]', '', x))

    df.iloc[:, 0] = df[0].str.strip()

    dct = df.groupby(0)[2].apply(list).to_dict()

    return dct

  else:

    df = df.iloc[:, 2]

    df = df.dropna()

    dct = {

       'standards': df.tolist()

    }

    return dct
```

```
# function to extract financial standards

def extract_standards(dct):

  dct_clean = {}

  prompt = '''

  This list contains standards for financial literacy.

  However, a single standard may be divided into multiple elements,

  or multiple standards may be combined within a single element.

  Please extract all standards accurately and return them as a list of strings.

  '''



  for key, lst_standards in dct.items():

    response = client.chat.completions.create(

      model="gpt-4o-mini",

      messages=[

        {

          "role": "system",

          "content": dedent(prompt)
```

```
    },
    {
        "role": "user",
        "content": str(lst_standards)
    }
],
response_format = {
    "type": "json_schema",
    "json_schema": {
        "name": "standards",
        "schema": {
            "type": "object",
            "properties": {
                "standards":{
                    "type": "array",
                    "items": {"type": "string"}
                }
```

```
                },
                "required": ["standards"],
                "additionalProperties": False
            },
            'strict': True
        }
    }
)
    str_response = response.choices[0].message.content
    dct_clean[key] = json.loads(str_response)['standards']


  return dct_clean


# function to clean noise from text
def clean_text(text):
  if pd.isna(text) or text == 'None':
```

```
    return ""


# Convert to lowercase

text = text.lower()



# Remove punctuation except underscores

text = re.sub(r"[,\-&$@#*+]", " ", text)

text = text.replace('’', "'")

text = re.sub(r"[^a-z_.?!' ]", '', text)

text = re.sub(r'\s+', ' ', text)



# Replace financial terms with underscore-connected versions

for term in FINANCIAL_TERMS:

    pattern = r'\b' + re.escape(term) + r'\b'
```

```
        replacement = term.replace(' ', '_')
        text = re.sub(pattern, replacement, text)

    return text

# function to apply preprocessing for NLP
def preprocess_text(text):
    text = clean_text(text)

    if text == "":
        return []

    # Tokenize
    tokens = word_tokenize(text)
```

```
    # Remove stop words
    stop_words = set(stopwords.words('english'))
    tokens = [token for token in tokens if token not in stop_words]


    # Lemmatize
    lemmatizer = WordNetLemmatizer()
    tokens = [lemmatizer.lemmatize(token) for token in tokens]


    return tokens


# function to apply all the above steps to the pdf
def preprocess_pdf(pdf_path):
    filename = os.path.basename(pdf_path)
```

```
filename_json = filename.replace('.pdf', '.json')

is_national = 'national' in filename




df = extract_table(pdf_path)

dct = df_to_dict(df, is_national)

dct = extract_standards(dct)




with open(f"./json/{filename_json}", 'w') as f:

    json.dump(dct, f, indent=4, ensure_ascii=False)




with open(f"./json/{filename_json}", 'r') as f:

    data = json.load(f)
```

```
    for key, lst in data.items():
        for i in range(len(lst)):
            lst[i] = ' '.join(preprocess_text(lst[i]))



    with open(f"./json_clean/{filename_json}", 'w') as f:
        json.dump(data, f, indent=4)


# function to create a batch file for openai embedding request
def create_batch_file(json_folder_path):
  lst_files = os.listdir(json_folder_path)
  batch_file = []
  for fname in lst_files:
    with open(json_folder_path + fname, 'r') as f:
      data = json.load(f)
      for key, value in data.items():
        for i in range(len(value)):
```

```
if fname == 'CEE_national.json':
    cid = f"CEE_national-{key}-{i}"
else:
    cid = f"{fname[:fname.index('-')]}-{key}-{i}"



batch_req = {
    "custom_id": cid,
    "method": "POST",
    "url": "/v1/embeddings",
    "body": {
        "input": value[i],
        "model": "text-embedding-ada-002",
        "encoding_format": "float"


    }
```

```
                }

                batch_file.append(batch_req)

    with open('./batch_request.jsonl', 'w') as f:

        for batch in batch_file:

            f.write(json.dumps(batch) + '\n')

!rm -r ./pdfs/.ipynb_checkpoints

lst_pdfs = os.listdir('./pdfs/')

# apply preprocessing to the documents

for pdf in lst_pdfs:

    preprocess_pdf(f'./pdfs/{pdf}')



create_batch_file('./json_clean/')

batch_input_file = client.files.create(

    file=open("batch_request.jsonl", "rb"),

    purpose="batch"

)
```

```
print(batch_input_file)

batch_input_file_id = batch_input_file.id

batch_request = client.batches.create(
  input_file_id=batch_input_file_id,
  endpoint="/v1/embeddings",
  completion_window="24h",
  metadata={
    "description": "paragraph embeddings"
  }
)




print(batch_request)

status = client.batches.retrieve(batch_request.id)

print(status)

file_response = client.files.content(status.output_file_id)
```

```
with open('response.json', 'wb') as f:
  f.write(file_response.content)
dct_embeddings = {}
with open('response.json', 'r') as f:
  for line in f:
    data = json.loads(line)
    cid = data['custom_id']
    fname, key, i = cid.split('-')


    if 'national' in fname:
      fname = f'{fname}.json'
    else:
      fname = f'{fname}-curriculum-alignment.json'
```

```
        if fname not in dct_embeddings:
            dct_embeddings[fname] = {}
        if key not in dct_embeddings[fname]:
            dct_embeddings[fname][key] = {}

        dct_embeddings[fname][key][i] = data['response']['body']['data'][0]['embedding']
for fname, response_content in dct_embeddings.items():
    data = {}
    with open(f'./json/{fname}', 'r') as f:
        data = json.load(f)
        for key, lst_standards in data.items():
            for i in range(len(lst_standards)):
                data[key][i] = [lst_standards[i], response_content[key][str(i)]]

    with open(f'./json_embeddings/{fname}', 'w') as f:
        json.dump(data, f, indent=4)
```

 # create a zip file of the output folder

!zip -r json.zip ./json

!zip -r json_clean.zip ./json_clean

!zip -r json_embeddings.zip ./json_embeddings

Appendix G - Domestic Policy Change

With these findings, I'm actively supporting curriculum development conversations with key figures in financial education. Tim Ranzetta, the co-founder of Next Generation Personal Finance, told me that I could participate in his weekly Thursday guest speaker series. This opportunity will let me inform thousands of educators about my findings into the overlooked learning objectives in financial education, so even if their state does not mandate it, they can still teach the material. Additionally, policy conversation with Joel Chrisler will let me work closer with the Jump$tart coalition to improve financial education nationwide, as they are a close partner with the CEE. The Jump$tart coalition co-authored the CEE national standards for financial education.

Appendix H - Personal Use of Findings

In my high school I run a personal financial literacy club that has reached over 200 students across 10 events. In the club's workshops or events, I specifically ensure that the club officers discuss the overlooked learning objective in classroom discussions. This process stems from the understanding that students need a comprehensive understanding of financial literacy for behavior change, so our club aims to not overlook the learning objectives in Appendix A.