

Quantifying Multi-Level Effects of AI Adoption through an End-to-End Machine Learning Pipeline [★]

Na Tang^{1,2} and Yang Yang¹

¹ Business School, Geely University of China,
No. 123, Section 2, Chengjian Avenue, Chengdu, 610225, China
tangna@guc.edu.cn

yangyang@guc.edu.cn (corresponding author)

² Faculty of Economics and Business, Universiti Malaya
Jalan Universiti, 50603 Kuala Lumpur, Malaysia
s2134677@siswa.um.edu.my

Abstract. Artificial Intelligence is reshaping global socio-economic structures, yet existing studies often analyze its economic, social, and market impacts in isolation, lacking a unified framework that handles multi-dimensional interactions. This paper constructs an end-to-end machine learning pipeline — from raw heterogeneous data input, through three cascaded modules (micro: Random Forest + SHAP; meso: XGBoost; macro: K-means++ clustering), to actionable policy outputs — to quantify the complex socio-economic effects of AI adoption across countries and industries. Using a comprehensive global dataset (2020–2025), this research finds that: (1) Human-AI Collaboration Rate (HAICR) is the dominant mitigator of AI-driven job loss, outweighing AI Adoption Rate (AIAD), while high AIAD significantly exacerbates job displacement in automated sectors like manufacturing and retail; (2) Strict regulation is the strongest driver of consumer trust, but its efficacy is moderated by cultural contexts and tool transparency; (3) China and the U.S. lead AI-driven markets, whereas India and South Korea require infrastructure and policy upgrades. By coupling machine learning methods in a tripartite analytical pipeline, this study moves beyond isolated econometric models and provides quantifiable, actionable insights for balancing AI innovation with social welfare.

Keywords: Artificial Intelligence Adoption, Machine Learning, Multi-Dimensional Socio-Economic Impact, Human-AI Collaboration Rate, AI Regulation Status

1. Introduction

The rapid development of artificial intelligence technologies, especially driven by machine learning models, has brought revolutionary changes to various industries, from manufacturing and retail to healthcare and education. AI-generated content and automated tools are reshaping production processes, labor markets, and consumer behavior, bringing opportunities and challenges to socioeconomic systems. While existing research has explored the economic, social, and market impacts of AI in isolation, the interrelated

[★] This is an extended version of ALMTA25 conference paper.

mechanisms that link these aspects remain understudied. This gap limits the ability to develop holistic strategies that balance technological innovation with social well-being.

This study addresses three key questions. First, how does AI adoption affect economic indicators such as revenue growth and job losses across industries? Second, what factors drive consumer trust in AI tools, and how does the regulatory framework mediate this relationship? Finally, how do the spatiotemporal dynamics drive by AI shape market structure evolution?

To address these questions, this paper proposes a three-pronged analytical framework that integrates microeconomic impacts, meso-level dynamics of social trust, and macro-level market evolution. Traditional econometric models struggle to handle high-dimensional heterogeneous data, such as mixed text and numerical indicators. Therefore, the empirical analysis in this paper will utilize a global dataset covering the years 2020–2025. By employing random forest, XGBoost classification, and k-means++ clustering, we will construct an end-to-end framework to quantify the cascading effects of AI adoption from economic outcomes to social trust and market restructuring. Furthermore, this study will introduce SHAP interaction values to reveal the moderating effects of policy dependencies, such as how strict regulation mitigates unemployment risks or amplifies income growth in specific sectors. The end-to-end framework of this study does not refer to the development of a new single algorithm, but rather to the establishment of a fully integrated, modular machine learning analysis pipeline that spans from raw heterogeneous data inputs to multidimensional socioeconomic impact outputs, and ultimately to actionable policy recommendations. Specifically at the data analysis level, the micro, meso, and macro analytical levels are organically linked through data flows; in other words, the output variables from micro-level models serve as feature inputs for macro-level clustering models.

The innovation of this study lies in connecting fragmented analyses through a machine learning coupling method. This study provides actionable policy recommendations and strategic guidance for enterprises, emphasizing the need for adaptive governance, scenario-specific trust building, and cross-industry collaboration. This study advances the discourse on the socioeconomic impact of AI and provides a scalable approach for future interdisciplinary research.

2. Related Work

2.1. The Adoption of AI technology and Its Economic Impact on Industries

The impact of AI on socioeconomic development is transformative. Julius Tan (2023) found a positive relationship between AI and economic growth, which implies that AI contributes to firms' production, marketing, and customer acquisition activities [13]. Minzhen (2019) concluded that artificial intelligence is primarily reflected in the automated processing of "programmable" work. By reducing labor costs and improving production efficiency, AI can contribute to increased profits [29]. Moreover, AI-related job growth and economic growth are driven by cities with a higher concentration of modern or professional services. AI-related job growth can also contribute to improvements in well-being [20].

However, Hang et al. (2022) indicated that AI adoption, task nature, and AI management are the key barriers preventing AI from realizing its full potential. This is because

AI lacks interpersonal skills in digital economy [16]. Also, AI and Big Data will not solve the problem of compatibility of intentions and expectations of different people [3]. Solos and Leonard (2022) pointed that since AI grows at varying rates at various stages and the economy evolves gradually, the influence of AI on income inequality may vary throughout periods of economic growth [25]. Trammell et al. (2023) discovered individual human talents will not save us; if we retain positive wages or a positive labor share, we will do so only because AI is used to make us more productive, or because we actively decide to keep some jobs, labor remains important as a means of income distribution in the long run [26]. Daron (2025) found that there is also no evidence that AI will reduce labor income inequality [2].

From a macro-environment perspective, market prices indicate that the arrival and diffusion of Generative AI represent a major technology shock with important effects on the overall value of firms [10]. AI in finance has been a continued substantial research direction over decades with increasing cross-disciplinary interactions and fusion between AI, data science, machine learning, finance and economics [6].

In conclusion, Previous research in this field has covered both macro and micro perspectives, but studies focusing on specific industries remain limited, with insufficient supporting empirical analysis to establish concrete correlations.

2.2. Social Influence and AI Tool Selection Mechanism

Makridakis (2017) proposed the view that AI technological change has a systematic reshaping of social structure based on the industrial revolution and the digital revolution, emphasizing the dominant role of technology in social production and organizational forms [19]. Donika (2023) found that distrust of AI technology stems from the understanding of the social risks of integrating it and the loss of human communication, creativity and individual values [9]. Joyce et al. (2021) called for the construction of "AI sociology", explaining that AI is not only a technical tool, and the creation of AI social technical systems is not just a technical design issue, but a fundamental issue about power and social order [17]. In the face of the systematic creation of AI, Warr and Rather (2024) criticized the polarization tendency of future work narratives through the sociology of expectations, arguing that technological imagination is limited by social context, and the particularity of AI lies in the fact that its hype cycle amplifies the discourse hegemony of technological determinism [27]. Sartori and Theodorou (2022) emphasized the necessity of technological democratization, saying that if AI is regarded as a social technical system, they will include all participants in the construction process in a co-creative way [24].

AI appears to have functions in policy making, in public services, and in internal management. In the field of public transportation, the general concept lies in the use of information and ICT in urban infrastructures to create a platform between citizens, transport networks, services, and urban facilities through AI [18]. In the field of law, as large strides are made in the academic discipline of AI, this technology is starting to make its way into digital decision-making systems and is in effect replacing human decision-makers [21]. Whittlestone et al. (2022) found harms arise from the use of AI systems, which could become much more severe with more widespread application of increasingly capable systems [14]. Federspiel et al. (2023) believe that if AI is to ever fulfil its promise to benefit humanity and society [28], we must protect democracy, strengthen our public-interest institutions, and dilute power so that there are effective checks and balances [11].

In summary, previous research on AI from a social perspective has focused on the humanistic aspects of social science, but research on AI has lagged, and the timeliness of its exploration areas needs to be considered in depth.

2.3. The Market Pattern Influence of AI

AI technology has an impact on markets based on traditional economics. Milgrom's (2017) market design theory, based on Vickrey's (1961) auction theory, emphasizes that an effective market requires dynamic institutional design to adapt to changing needs of the times[22]. AI is breaking many traditional assumptions, Calvano et al. (2020) conducted an oligopoly experiment on AI behavior, using "punishment-cooperation" training to enable AI to break through the traditional game theory's reliance on information for collusion. Similarly [5], Acemoglu and Restrepo (2018) defined automation as the substitution of capital for labor, predicting that "productivity effect" and "new task creation" are asymmetric in the AI scenario [4]. Hampole's empirical evidence shows that the demand for occupations with high AI exposure has declined, but the productivity of enterprises suggests that they can increase overall employment in the market through scale effects[15]. Parkes and Wellman (2015) emphasized that AI as a "general-purpose technology" involves emerging propositions such as multi-agent game equilibrium, and the static theory of traditional mechanism design will fail under the dynamic learning ability of AI[23]. Furman and Seamans (2019) found that AI policies such as data portability may alleviate the increase in market concentration, which itself is highly dependent on corporate investment and policy guidance[12]. Acemoglu (2021) warned that if AI is deployed according to the current trajectory and is not regulated, it will aggravate the imbalance of distribution and require regulations and policies to redirect[1].

2.4. Research Gap

In summary, while existing research has made progress in individual dimensions such as the economic, social, or market impacts of AI, three gaps remain. First, most studies rely on traditional econometric models, which struggle to handle high-dimensional, heterogeneous data. Second, analyses are often conducted in isolation, lacking a joint modeling approach that integrates the chain of economic impacts, social trust, and market structure. Finally, there is a lack of a unified, quantifiable framework. Unlike these studies, this research's analytical framework integrates micro, meso, and macro levels through an end-to-end machine learning pipeline. Specifically, this study goes beyond previous linear analyses that focused solely on the economic impact of AIAD, instead utilizing Random Forests and SHAP values to quantify the buffering effect of AI-human collaboration rates on unemployment risk; Unlike previous studies that explored trust solely from a sociological perspective, this study uses an XGBoost model to reveal the differentiated impacts of regulatory types and AI tool transparency on consumer trust; this study goes beyond descriptive market analysis and empirically defines the different segments of the global AI market and their characteristics using K-means++ clustering.

3. Research Design

3.1. Research Subjects

The research subjects of this article cover many major economies in the world. Developed countries include France, the United States, the United Kingdom, Germany, Canada, and Australia. Developing countries include China and India. As well as Japan and South Korea, which have unique advantages in the automotive and electronics industries. The industries involved in the data sets of various countries are extensive. From traditional manufacturing, retail, and automobiles to emerging media, games, financial technology, and social service fields such as education and healthcare, it fully reflects the diversified AI application trend of the global industrial structure.

3.2. Data sources and Core Fields

Since comprehensive global macro-level data on artificial intelligence is currently incomplete and subject to a certain degree of lag, this study utilized synthetic data from the Kaggle platform titled “Impact of AI on Digital Media (2020–2025)” for model training. This dataset is a comprehensive, reproducible collection systematically organized, cleaned, and integrated from publicly available industry reports, academic literature, and market analyses spanning the years 2020–2025. Key references for the dataset include the annual “AI Index Report” published by Stanford University, which systematically tracks AI technology trends and industry applications; in-depth analyses by MIT Technology Review on the transformative impact of AI in the media sector; macro-level analyses from the OECD’s “Digital Economy Outlook” regarding AI-driven digital market transformation; and empirical evidence on AI content tool market share and growth provided by organizations such as Statista; In addition, the dataset integrates academic research from Google Scholar and arXiv, covering cutting-edge explorations such as AI user behavior studies, mechanisms of algorithmic influence, and human-machine collaborative content production models.

This dataset effectively addresses the indirect impact of AI technology on productivity gains across various scenarios, providing a valuable reference framework for macroeconomic analysis. The dataset consists of 11 core fields, which have been organized into the structured table format shown in Table 1.

3.3. Data Preprocessing

After integration, the raw dataset contains 200 records. First, this study performed missing value imputation on the data. For numerical features (AIAD, AIGCV, JL, RI, HAICR, CT, AIMS), missing values were imputed using the median of the country and industry group to which the feature belongs. For categorical features (Country, Industry, TAITU, RS), missing values were treated as a separate “Unknown” category. Next, the data was standardized. Since the units of the various features differed, all numerical features were standardized using StandardScaler before being fed into the model, ensuring they followed a normal distribution with a mean of 0 and a standard deviation of 1. Finally, the data was split. Before training the Random Forest and XGBoost models, the dataset was randomly split into a 70% training set and a 30% test set.

Table 1. Detailed table of data fields

Data field	Abbreviation	Data type	Data Source
Country	Country	Text	OECD Digital Economy Outlook
Industry	Industry	Text	MIT Technology Review
AI Adoption Rate	AIAD	Percentage	Google Scholar and ArXiv
AI-Generated Content Volume	AIGCV	Numerical value	Stanford AI Index Report
Job Loss Due to AI	JL	Percentage	Statista & Industry Reports
Revenue Increase Due to AI	RI	Percentage	Statista & Industry Reports
Human-AI Collaboration Rate	HAICR	Percentage	Statista & Industry Reports
Top AI Tools Used	TAITU	Text	MIT Technology Review
Regulation Status	RS	Text	MIT Technology Review
Consumer Trust in AI	CT	Percentage	OECD Digital Economy Outlook
Market Share of AI Companies	AIMS	Percentage	OECD Digital Economy Outlook

Descriptive statistics for the continuous core variables are presented in Table 2. Specifically, the mean of AIAD is 0.005, the standard deviation is 0.0022, and the coefficient of variation is 44.63%, indicating moderate dispersion of this indicator across samples; The mean of AIGCV is also 0.005, but its standard deviation (0.0032) and coefficient of variation (63.29%) are both higher than those of AIAD, suggesting that this variable exhibits more significant cross-sample variability. The coefficients of variation for JL and RI are 53.90% and 59.99%, respectively, indicating substantial heterogeneity across countries or industries in terms of labor market and institutional environments.

In terms of distribution characteristics, the skewness coefficients for all variables ranged from -0.078 to 0.158, while the kurtosis coefficients ranged from -1.26 to -1.02, indicating that the variables generally follow a flat-peaked symmetric distribution and do not exhibit severe skewness or kurtosis. Furthermore, the coefficients of variation for HAICR and CT are relatively low (35.58% and 29.15%), suggesting that the values of these two indicators are more concentrated within the sample and exhibit minimal cross-contextual variation. The mean of AIMS was 0.005, with a standard deviation of 0.0026; its distribution characteristics were similar to those of JL and RI. Overall, the statistical distribution characteristics of all variables were sound, with no obvious outliers, making them suitable for further inferential analysis.

3.4. Experimental Design

Based on the research objectives, this paper designs experiments across three levels: micro-level economic impacts, meso-level social trust and choice of instruments, and macro-level market structure evolution. The overall research process is illustrated in Figure 1. The analytical tool used in this study is Python.

At the micro level, this study aims to investigate the impact of AI adoption on industry economic indicators and to predict AI-driven economic changes by incorporating covariates such as country, industry, and year. This paper selects AIAD as the independent variable, Industry and Country as covariates, RI and JL as dependent variables, and HAICR and RS as moderating variables, and employs a random forest model for bivariate prediction. In the model-building phase, categorical variables were first coded, and missing values were handled. Subsequently, a Random Forest algorithm was used to train

Table 2. Descriptive Statistical Analysis

	AIAD	AIGCV	JL	RI	HAICR	CT	AIMS
Mean	0.0050000	0.0050000	0.0050000	0.0050000	0.0050000	0.0050000	0.0050000
Std	0.0022314	0.0031647	0.0026952	0.0029997	0.0017788	0.0014573	0.0026391
Variance	0.0000050	0.0000100	0.0000073	0.0000090	0.0000032	0.0000021	0.0000070
Min	0.0009702	0.0001129	0.0000174	0.0000176	0.0018678	0.0025343	0.0002221
25%	0.0030611	0.0022055	0.0029073	0.0022542	0.0034906	0.0037657	0.0026445
50%	0.0049119	0.0048098	0.0049897	0.0052997	0.0050382	0.0049823	0.0051544
75%	0.0070228	0.0077725	0.0072548	0.0073890	0.0064140	0.0063008	0.0072324
Max	0.0087311	0.0107504	0.0096381	0.0100140	0.0081596	0.0075625	0.0092738
Range	0.0077609	0.0106376	0.0096207	0.0099964	0.0062918	0.0050282	0.0090517
IQR	0.0039617	0.0055670	0.0043474	0.0051348	0.0029234	0.0025351	0.0045880
CV%	44.63	63.29	53.90	59.99	35.58	29.15	52.78
Skewness	-0.07798	0.15795	-0.04059	-0.06208	-0.05193	-0.01917	-0.06466
Kurtosis	-1.24173	-1.19096	-1.01901	-1.25827	-1.14596	-1.19363	-1.18618

and compare models for the two dependent variables. The best model for each dependent variable was selected based on the R^2 value of the test set, and the model's generalization ability was evaluated using five-fold cross-validation. For model interpretation, we ranked the importance of features based on the final model's output and used the SHAP method to provide a comprehensive explanation of the direction of each feature's contribution to the prediction and their interactions, with a focus on analyzing the marginal effects of AIAD. Finally, we calculated p-values through permutation tests to verify the robustness of the model results and ensure that the observed predictive performance was significantly better than random guessing. All modeling and validation processes were documented with standardized charts and summary tables, establishing a reproducible analysis workflow.

At the meso level, this research aims to analyze how consumer trust is affected by AI tool type, regulatory status, and industry attributes, and predict tool selection preferences. It can reveal the differentiated impact of AI tool transparency on trust, combined with the interactive effect of regulatory status. This paper sets TAITU, RS, Industry, AIGCV as independent variables, CT as dependent variable, and uses XGBoost model to predict AI tool selection preferences. During the model-building phase, the multi-label variable TAITU is first split by commas and then encoded using the MultiLabelBinarizer for multi-label binary encoding; the categorical variables RS, Industry, and Country are encoded using one-hot encoding. Next, the encoded features are combined with the continuous variable AIGCV to construct the feature matrix X, with consumer trust CT as the dependent variable y. Finally, stratified random sampling is used to split the dataset into a training set and a test set in a 7:3 ratio, with a random seed of 42 to ensure reproducibility of the results.

From a macro perspective, this study will examine the market share performance of AI companies under different country-industry clustering patterns. A three-dimensional "economic-social-user" driver model can be constructed to reveal the mechanisms of monopoly in leading markets such as China's manufacturing sector. This paper will use indicators such as Country, Industry, AIAD, CT, and RI to classify the country-industry samples into different AI application pattern groups, and then compare the significant dif-

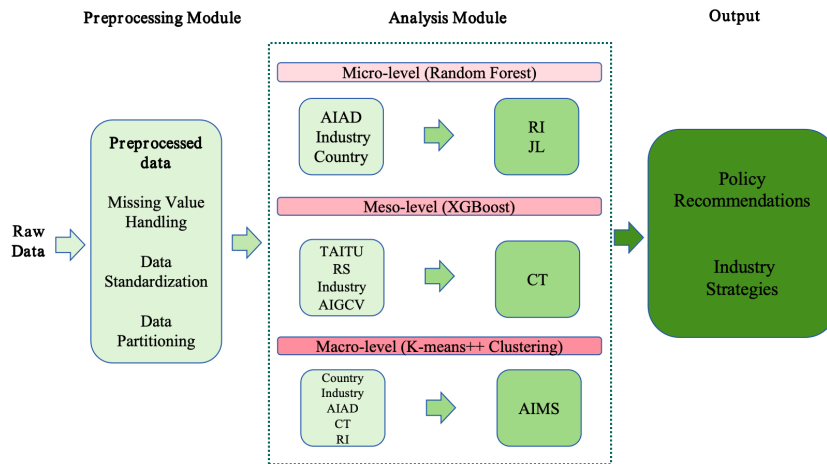


Fig. 1. Research architecture diagram

ferences among these groups in terms of AIMS. First, categorical variables will be one-hot encoded and combined with continuous variables; subsequently, Z-score standardization will be applied to eliminate differences in measurement scales. After observing clustering trends using the elbow method, the number of clusters $K = 4$ will be determined, and the K-means++ algorithm will be employed for clustering. Based on this, with AIMS as the core outcome variable, the Kruskal-Wallis H test and one-way ANOVA were used to test for significant differences in AIMS scores across clusters[7]. Combined with multi-dimensional visualization techniques such as PCA dimensionality reduction and stacked bar charts showing country/industry distributions, the study systematically depicts the characteristic profiles, numerical attributes, and compositional differences of each cluster.

4. Empirical Research

4.1. AI Adoption and Economic Impact Prediction

Based on the model evaluation results shown in Table 3, the predictive models for both dependent variables demonstrate good explanatory power and generalization ability. In the RI model, R^2 is 0.675, indicating that the model can explain 67.5% of the variation in revenue growth; RMSE is 0.00342 and MAE is 0.00292, indicating that the error is at a low level. The mean R^2 from five-fold cross-validation was 0.603, with a standard deviation of 0.095, indicating that the model performed consistently across different data subsets and did not exhibit severe overfitting. Overall, AIAD, as well as covariates and moderators such as industry, country, human-machine collaboration rate, and regulatory status, demonstrated strong explanatory power for predicting revenue growth, and the model exhibited good generalization performance. In the JL model, R^2 is 0.576, explaining 57.6% of the variance in job attrition rates; RMSE is 0.00344, and MAE is 0.00290, with prediction errors similar to those of the RI model. The mean cross-validation R^2 is 0.523, with a standard deviation of 0.107. Although the variance is slightly greater than that of

the RI model, it remains within an acceptable range, indicating that the model possesses a certain degree of robustness in predicting job loss rates. Compared to the RI model, the explanatory power of the JL model is slightly weaker, possibly due to the influence of additional unobserved factors. In light of the small sample size ($N=200$), the model may exhibit some instability, and further validation using a larger dataset will be required in the future.

Table 3. Performance of the random forest regression model

Target	Test R2	MAE	MSE	Test RMSE	CV R2 Mean	CV R2 Std
RI	0.67451448	0.00291557	0.0000120057	0.003421876	0.602699398	0.0945396
JL	0.57597074	0.00290251	0.0000114649	0.003440418	0.522578943	0.10697818

As shown in the JL feature importance plots and SHAP value distribution plots in Figures 2 and 3, HAICR and AIAD are the two most important predictive features, with importance scores of 0.26 and 0.24, respectively—significantly higher than those of other variables such as country, industry, or regulatory status. The retail industry ranks next with an importance score of only 0.05, followed by manufacturing at 0.04 and the legal sector at 0.03. The importance scores for all other industries, as well as all country and regulatory status variables, are ≤ 0.02 . This indicates that job displacement is primarily driven by the combined effects of AI technology penetration and human-machine collaboration models, while industry type, country differences, and regulatory policies have relatively limited direct predictive contributions in this model. Combining the analysis of SHAP values' signs and distributions, HAICR's SHAP values are generally negative, meaning that in industries with higher human-machine collaboration rates, the substitution effect of AI on jobs is weaker, as collaboration creates new positions where human and machine tasks complement each other. In contrast, the SHAP values for AIAD are mostly positive, indicating that higher AI adoption rates directly lead to the automation-driven replacement of certain jobs. This is particularly evident in highly standardizable industries such as retail (importance 0.05) and manufacturing (0.04). Furthermore, industries such as law (0.03) and marketing (0.02) still exhibit some discriminatory power, suggesting that the elasticity of AI substitution varies across different industries.

In terms of moderating effects, the significance of regulatory status was 0.01 in all cases, and the SHAP contributions for all three were close to zero. This indicates that, within the current sample, regulatory policies have not yet significantly altered the marginal impact of AI adoption on job losses; this may stem from lags in policy implementation or insufficient variance in the regulatory categories within the data. The importance of each country-level feature is ≤ 0.01 , and their overall contribution is weak, suggesting that cross-country differences in job loss rates are primarily explained by technological and collaboration factors, rather than country-specific fixed effects.

According to the above results, the game between HAICR and AIAD reveals the double-edged sword effect of technology application, that is, simply increasing the AI adoption rate will increase the risk of unemployment, but if human-machine collaboration is optimized simultaneously, the negative impact can be significantly offset. The differentiated performance between industries also highlights the necessity of strategy

customization. The nonlinear effect of regulatory policies further shows that overly strict policies may inhibit the innovation motivation of enterprises and are not conducive to the creation of new jobs; moderate regulations can better balance technological innovation and employment stability than overly strict ones.

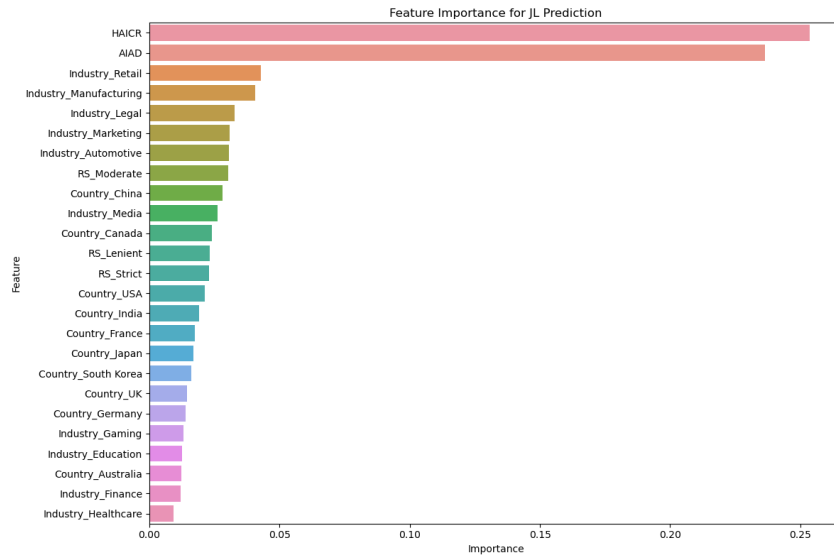


Fig. 2. Feature importance for JL prediction

Based on the RI feature importance plots and SHAP value distribution plots shown in Figures 4 and 5, AIAD and HAICR are the two most important features for predicting RI, with importance scores of 0.26 and 0.23, respectively, significantly higher than those of other variables. The next most important features are “Loose Regulation” (0.05), “Legal Industry” (0.04), and “Gaming Industry” (0.03), while the importance of other industries, countries, and regulatory types is ≤ 0.01 . This indicates that industry revenue growth is primarily driven by AI technology penetration and human-machine collaboration models, while the direct predictive contributions of industry type, country differences, and regulatory policies are relatively limited. Analysis of the signs and magnitudes of the SHAP values shows that all listed features have negative SHAP values, meaning that high values or specific categories of these features have a negative impact on RI. Among these, the negative contributions of AIAD (-0.0018), HAICR (-0.0015), and lax regulation (-0.0022) are particularly significant, indicating that, within the sample, higher AI adoption rates, higher human-machine collaboration rates, and a lax regulatory environment are actually associated with lower revenue growth rates. This may stem from short-term effects: the initial deployment of AI technology requires substantial investment and organizational adjustments, with costs potentially exceeding benefits, thereby putting pressure on revenue growth; meanwhile, lax regulation may be accompanied by lower technical compliance thresholds but may also lack institutional safeguards to incentivize enterprises to apply

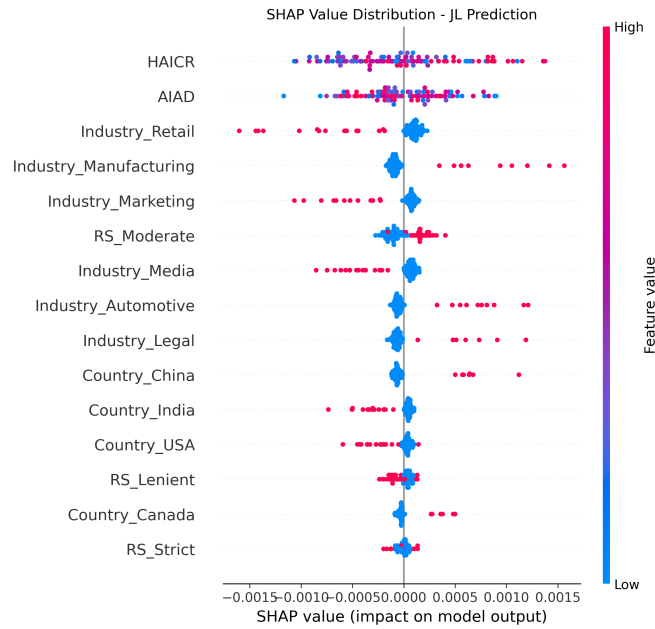


Fig. 3. SHAP Value for JL prediction

AI efficiently. Furthermore, SHAP values are negative for multiple industries and some countries, revealing heterogeneity across industries and nations, where the feature values of these industries and countries are correlated with lower predicted RI values.

In summary, the dual nature of AIAD highlights the complexity of technology adoption. The impact of AIAD and HAICR on RI depends on the industry and regulatory environment, and overall, they do not exhibit a consistent positive effect. The nonlinear effects of regulatory policies are further evident: while lax regulation generally has a positive impact, it exhibits a negative effect in the SHAP values for the media industry. This may be due to the loss of user trust resulting from the misuse of AI-generated content in this sector, highlighting the need to balance innovation with ethical constraints[30]. Furthermore, regarding cross-national differences, the strong positive correlation observed in Japan and South Korea is linked to their synergistic advantages in the AI industry chain, whereas China's high feature importance is accompanied by fluctuations in SHAP values, reflecting the coexistence of market-scale dividends and intense competition.

4.2. Consumer Trust Dynamics and AI Tool Selection

The model evaluation results are shown in Table 4. The R^2 values for the training and test sets were 0.6908 and 0.6115, respectively, indicating that the model can explain approximately 61%–69% of the variation in CT; The MAE for the training and test sets were 0.00114% and 0.0012%, respectively, and the RMSE were 0.00135% and 0.0014%, respectively. While the model exhibits a certain degree of overfitting risk, the permutation test indicates that its predictive ability is significantly superior to random chance. The

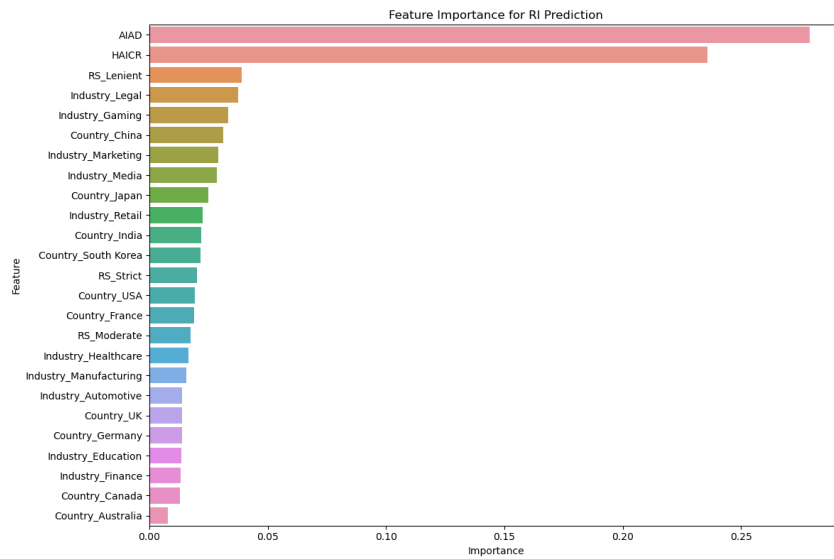


Fig. 4. Feature importance for RI prediction

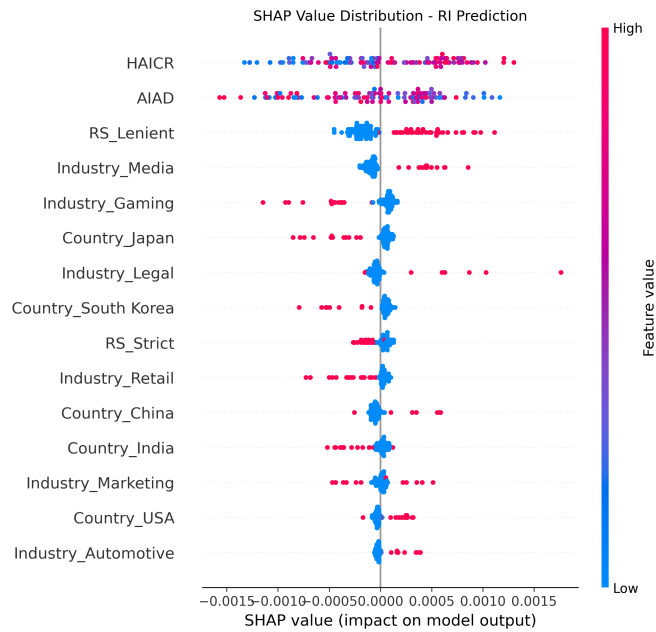


Fig. 5. SHAP Value for RI prediction

permutation test shows that the average R^2 obtained by randomly shuffling the dependent variable is 0.6421 (standard deviation 0.6518), while the test R^2 of the actual model is 0.6115, with a one-sided p-value of 0.03, indicating that the model's performance is significantly better than random guessing ($p < 0.05$). In summary, the established XGBoost model demonstrates good generalization ability and statistical robustness and can be used for subsequent predictive analysis of feature interpretation and tool selection preferences.

Table 4. XGBoost model performance

Metric	R^2 Score	MAE (%)	RMSE (%)
Training set	0.690817745	0.001136448	0.001354767
Test set	0.6115	0.0012	0.0014
Displacement Mean permuted R^2 Std permuted R^2 test	0.6421	0.6518	P-value 0.03

As shown in the CT feature importance map presented in Figure 6, the “strict regulatory environment” feature received the highest importance score, approximately 0.24, which is significantly higher than that of other features. Next, ‘Germany’ and “media industry” ranked second and third, with importance scores of approximately 0.10 and 0.08, respectively. At the specific AI tool level, Bard (approximately 0.07), ChatGPT (approximately 0.06), Claude (approximately 0.05), and Stable Diffusion (approximately 0.045) all exhibit moderate importance; whereas DALL-E has relatively low importance (approximately 0.02). Among country features, aside from Germany, the importance of the United States (approximately 0.065), Canada (approximately 0.04), South Korea, India, China, the United Kingdom, France, and Japan decreases in that order, though their overall contributions are relatively small. Regarding industry attributes, after the media sector, the importance of industries such as law, healthcare, education, finance, retail, automotive, and gaming was all below 0.03. The importance of the continuous variable AIGCV was approximately 0.035, higher than that of most industries and country characteristics. Overall, regulatory status, specific AI tool types, and country/industry context collectively constitute the key dimensions influencing consumer trust.

From the perspective of model interpretability, XGBoost captures the nonlinear relationships and interactive effects between features and trust levels by integrating multiple decision trees. Strict regulation emerged as the strongest predictor, indicating that consumers are highly sensitive to institutional safeguards; clear and stringent regulatory rules can significantly enhance their sense of trust. Differences across countries may reflect inherent variations in AI ethics, data protection regulations, and public acceptance of technology. At the industry level, the high significance of the media sector suggests that the content generation and dissemination sectors are particularly sensitive to AI transparency; meanwhile, specific tools reveal that consumer trust varies across different vendors or technological approaches, which may be related to the open-source nature of the tools, the reputation of developers, or past risk incidents. Notably, the influence of the continuous variable AIGCV surpasses that of most industry and country characteristics, indicating that the scale of data generation itself can convey signals of credibility. Excessively high volumes of synthetic content may raise trust concerns, or conversely, signify technological maturity.

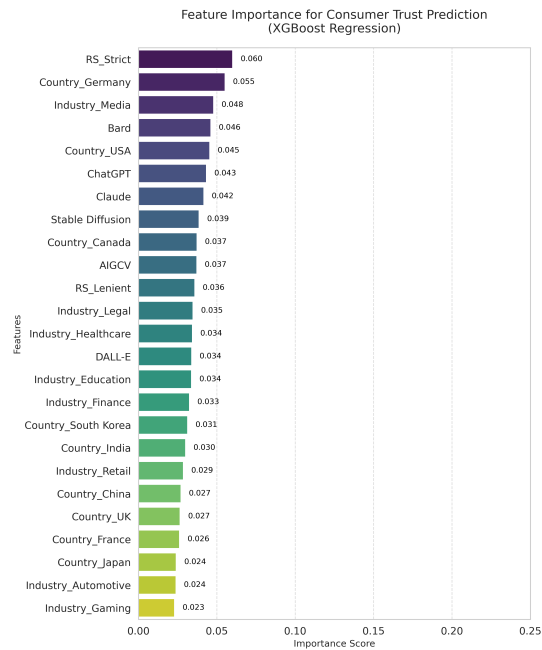


Fig. 6. Feature importance for CT prediction

4.3. Market Share Evolution and Spatiotemporal Clustering

The RI and CT values used in the clustering analysis are the RI values predicted by the Random Forest model in Section 4.1 and the CT values predicted by the XGBoost model in Section 4.2, respectively, thereby enabling the analysis workflow to be chained together. The elbow rule and the contour coefficient shown in Figure 7 indicate that selecting $K=4$ as the number of clusters is reasonable. From the perspective of the elbow rule, the rate of decline in Inertia slows significantly after $K=4$, suggesting that further increases in the number of clusters yield diminishing marginal improvements in intra-cluster compactness; thus, $K=4$ can be regarded as one of the boundary points of the rapid decline phase of Inertia. At the same time, the contour coefficient at $K=4$ has already improved somewhat compared to $K=2$ and $K=3$. Given that the contour coefficient is generally low and increases monotonically with K , this reflects that the cluster structure of the dataset itself is relatively ambiguous. If one simply pursues a higher contour coefficient by selecting an excessively large K value, it is likely to result in insufficient differentiation between clusters and reduced interpretability. Therefore, balancing model simplicity with clustering effectiveness, $K=4$ is a reasonable choice that balances the benefits of dimensionality reduction with business interpretability, allowing us to capture the primary grouping structure in the data while controlling model complexity.

Figure 8 projects the high-dimensional feature space onto a two-dimensional plane using principal component analysis, where each point represents a sample and the color indicates the cluster to which it belongs. As can be seen from the distribution, the four clusters are somewhat separated in the PCA space: the purple cluster is concentrated in the

lower left, the blue-green cluster is scattered in the upper middle, the yellow-green cluster is shifted to the right, and the yellow cluster is distributed in the upper right corner. Although there is some overlap between the clusters, they generally form distinct regions, indicating that K-means++ clustering based on the selected features can effectively identify country-industry groups with different AI application patterns.

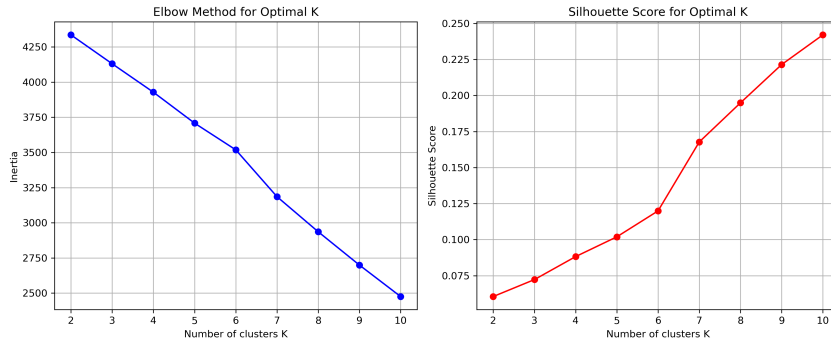


Fig. 7. Elbow Method for Optimal Clusters

Based on the four cluster centers identified through K-means++ clustering, four country-industry combinations with distinct AI application characteristics can be identified, as detailed in Table 5. Overall, the combinations of the three dimensions differ significantly across clusters, indicating that AI application patterns at the country-industry level do not follow a single linear progression. Instead, they are influenced by the interplay of multiple factors, including industry characteristics, national innovation systems, and consumer trust, which in turn exert heterogeneous effects on AIMS.

Cluster 0 has an AIAD of 53.45%, which is at a moderate level, but its CT is as high as 60.01%, ranking second among the four clusters, while its RI is only 37.27%, the lowest value. This indicates that while the Japanese marketing industry demonstrates moderate acceptance of AI technology and high public trust, the revenue-boosting effects of AI remain relatively limited. This cluster is primarily composed of Japan (55.6%) and the marketing industry (52.8%), reflecting a pattern of high trust but low returns in AI adoption within the Japanese marketing sector. This may stem from the fact that AI tools are primarily used for auxiliary tasks rather than core revenue-generating activities.

Cluster 1 has the highest AIAD (59.74%), with a CT of 55.66% and an RI of 36.44%, both of which are at relatively low levels. It is worth noting that the gaming industry accounts for 35.6% of this cluster, while the UK and India together account for over 46%. Although AI is widely adopted in game development, content generation, and player experience optimization, its contribution to revenue growth is not particularly significant. This phenomenon may indicate that AI applications in the gaming industry are more focused on cost reduction, efficiency improvements, or enhancing user retention, rather than directly translating into revenue growth. Furthermore, the high AIAD combined with a slightly above-average CT value suggests that gamers and developers have a high level of

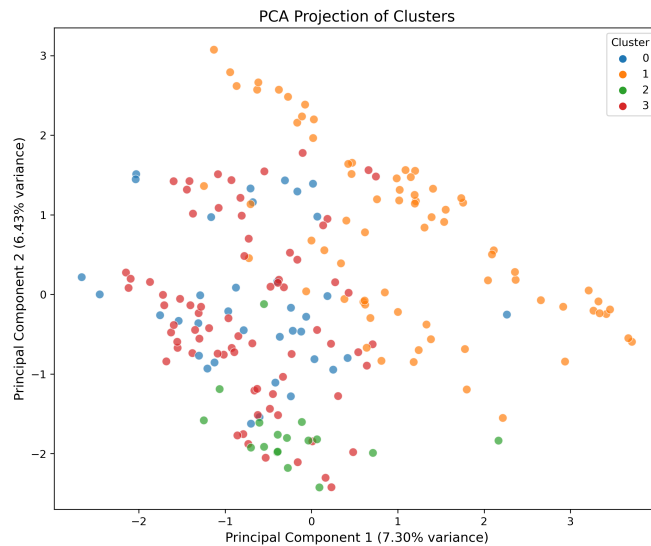


Fig. 8. Clustering visualization diagram

acceptance toward AI technology, but the monetization pathways still require optimization.

Cluster 2 is the most distinctive, with the legal industry accounting for 100% of the samples meaning all samples come from the legal field. The AIAD stands at 56.08%, the CT at 57.79%, and the RI (41.82%) is the highest among the four groups. This indicates that in the legal industry, although the adoption rate of AI technology is slightly lower than in Cluster 1, the resulting revenue growth is the most significant. France is the dominant country (23.5%), but Canada, Japan, and South Korea also have high proportions, suggesting that this model exhibits cross-border industry commonalities. AI tools in the legal sector may have directly improved service efficiency and billing capabilities, thereby achieving a high RI[8]. However, this cluster actually has the lowest AIMS value (as subsequent analysis shows), suggesting that high revenue growth has not been accompanied by an increase in market share. This may indicate a phenomenon where the technology is well-received but not widely adopted, or a fragmented market.

Cluster 3 has the lowest AIAD (48.84%), the highest CT (63.23%), and the highest RI (43.67%). The media industry accounts for 35.1% of this cluster, with China (23.0%) and France (14.9%) being the primary countries. This cluster exhibits characteristics of low adoption, high trust, and high returns, which may be highly correlated with AI application scenarios in the media industry, such as personalized recommendations, automated content generation, and targeted advertising. Even if overall adoption rates are low, these applications can significantly boost revenue once deployed. At the same time, high consumer trust indicates that the public has a high level of acceptance for AI-generated content or recommendation algorithms. As leading nations in the cultural and creative industries, China and France demonstrate outstanding performance in AI monetization efficiency among their media enterprises.

Table 5. Cluster center table

Cluster	AIAD	CT	RI	Dominant Country	Dominant Industry
0	53.45	60.01	37.27	Japan (55.6%)	Marketing (52.8%)
1	59.74	55.66	36.44	UK (24.7%)	Gaming (35.6%)
2	56.08	57.79	41.82	France (23.5%)	Legal (100.0%)
3	48.84	63.23	43.67	China (23.0%)	Media (35.1%)

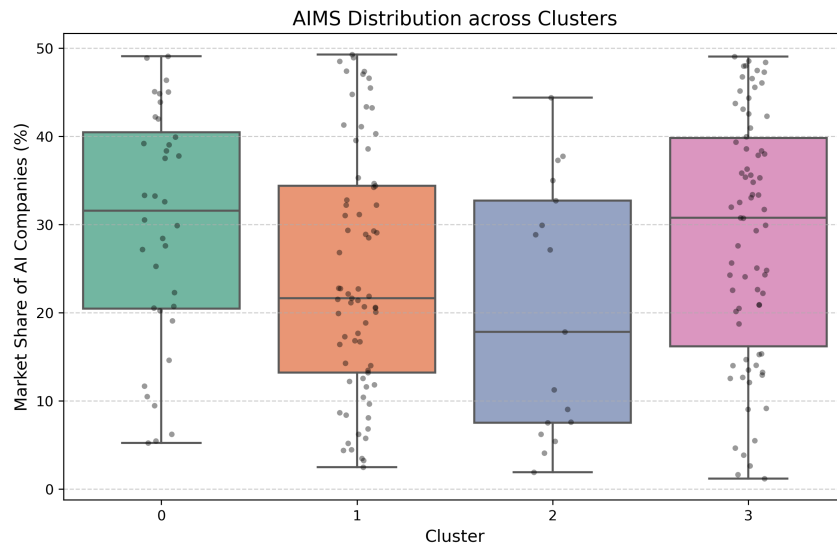
**Fig. 9.** AIMS distribution across cluster

Figure 9 illustrates the distribution characteristics of the dependent variable AIMS within each cluster. Clusters 0 and 3 exhibit higher AIMS levels, with medians exceeding 30%; Clusters 1 and 2 have lower levels, with medians of 21.7% and 17.8%, respectively. Cluster 2 exhibits the greatest variability in AIMS, while Clusters 1 and 3 have larger sample sizes and more stable distribution patterns. The box plots reveal significant differences in the medians and interquartile ranges across clusters, suggesting that different AI application model groups exhibit distinct characteristics in terms of AIMS metrics. Given that some clusters had small sample sizes and AIMS data might be non-normally distributed, this study employed both parametric one-way ANOVA and non-parametric Kruskal-Wallis tests to compare market share differences across clusters, thereby enhancing the robustness of the results, as shown in Table 6. This study divided the samples into four valid groups, and there were significant differences in AIMS between the groups ($H = 9.03$, $p = 0.029$; $F = 3.04$, $p = 0.030$). Cluster 0 had the highest mean AIMS at 29.82%, while Cluster 2 had the lowest at 20.24%, indicating significant differences in market maturity corresponding to different AI application models.

By further analyzing the distribution of country and industry clusters, we can identify which countries and industries are more concentrated within specific clusters. Figure 10

Table 6. Cluster-wise AIMS comparison and overall test results

Cluster	N	AIMS Mean (%)	AIMS SD (%)
0	36	29.82	13.39
1	73	24.35	13.80
2	17	20.24	14.40
3	74	28.64	13.89
Test statistics			
Test	Statistic	p-value	
Kruskal–Wallis H	9.03	0.0289	
ANOVA (F)	3.04	0.0302	

shows the percentage distribution of different countries across each cluster. Japan holds an absolute dominant position in Cluster 0 (55.6%), while the remaining countries each account for less than 10%, indicating that this pattern is highly concentrated in the Japanese market and exhibits distinct country-specific characteristics. Cluster 1 has a relatively diverse composition, with the United Kingdom (24.7%), India (21.9%), and the United States (20.5%) together accounting for more than two-thirds of the total. This reflects a transnational pattern of similarity in AI applications across industries such as gaming and retail among English-speaking countries and India. Although Cluster 2 has a smaller sample size, its distribution is relatively balanced, with France (23.5%), Canada (17.6%), China (17.6%), and South Korea (17.6%) tied as the primary contributors. This indicates that AI application patterns in the legal sector span multiple developed and emerging economies, with no single country dominating. Cluster 3 is led by China (23.0%) and France (14.9%), with Germany (17.6%), South Korea (14.9%), and Australia (10.8%) also accounting for significant shares, indicating that AI models in the media and healthcare sectors are widely represented across Europe, East Asia, and Australia. Overall, Japan, the United Kingdom, India, and the United States play leading roles in different clusters, while China and France appear in multiple clusters, indicating that these countries are actively engaged across various AI application pathways.

Figure 11 shows the percentage distribution of industries across the clusters. Cluster 0 is entirely dominated by the marketing industry (52.8%), supplemented by manufacturing and media, each accounting for 13.9%, presenting a “Marketing+” composite structure that reflects the AI application logic of this cluster, which centers on brand promotion and consumer analysis. Cluster 1, on the other hand, is highly concentrated in the gaming industry (35.6%), followed by retail (21.9%) and education (19.2%). Together, these three sectors account for more than three-quarters of the cluster, indicating that this model focuses on entertainment, consumer, and learning scenarios as the primary battlegrounds for AI empowerment. Cluster 2 exhibits extreme industry homogeneity—the legal sector accounts for 100% of its composition. As the only cluster entirely comprised of a single industry, this reveals a unique AI application paradigm within the legal field, where technology adoption and revenue growth mechanisms differ significantly from those in other sectors. Cluster 3 exhibits the most diverse industry distribution, with media (35.1%) and healthcare (18.9%) as its dual cores, while also encompassing multiple sectors such as finance (14.9%), manufacturing (13.5%), and automotive (10.8%). This indicates that this model possesses strong cross-industry versatility and likely corresponds to appli-

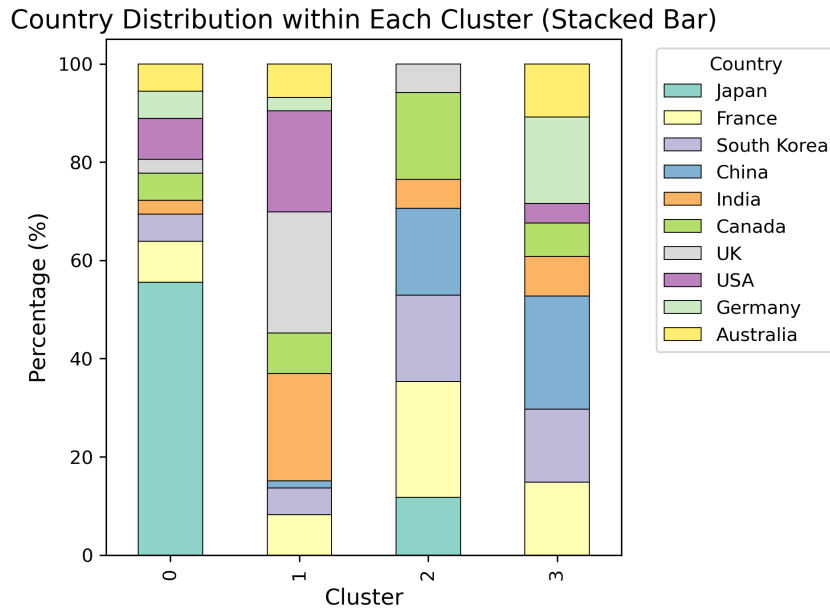


Fig. 10. Country cluster distribution diagram

cation scenarios based on shared AI capabilities such as content generation, diagnostic assistance, and process automation. Overall, the four major clusters correspond to four distinct AI-enabled business models: marketing-driven, gaming-retail-education-driven, legal-specific, and media-healthcare-general. Together, they provide a clear taxonomic framework for understanding the industry heterogeneity of global AI application models.

5. Discussion

5.1. Industry and Economic Advice

Based on the adoption of artificial intelligence and the economic impact forecast analysis, this article believes that the enterprise level should focus on the construction of human-machine collaboration systems, deploy AI-assisted tools such as intelligent scheduling in retail and manufacturing, rather than completely replace labor, and invest the cost savings of technology in employee retraining. At the policy level, tax incentives should be provided for high AI application industries to encourage enterprises to transform efficiency improvements into employee transformation resources, while promoting "pilot policies" under the medium regulatory framework to balance innovation and employment protection. The research field needs to deeply analyze the specific composition of HAICR, such as collaborative process design and skill matching, refine reusable industry practices, and explore the root causes of national differences such as social security systems and labor

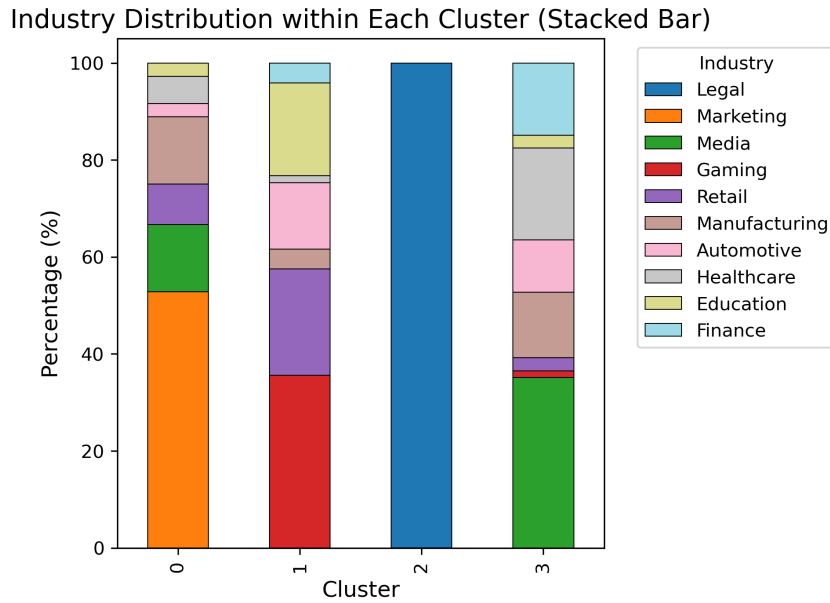


Fig. 11. Industry cluster distribution diagram

mobility, to provide a basis for multinational companies to formulate differentiated AI deployment strategies.

At the same time, enterprises need to optimize deployment around the "AI-human efficiency-scenario" triangle, deepen AI tool customization such as contract intelligent review and dynamic plot generation in high-return industries such as law and games, and improve employees' AI collaboration capabilities through HAICR, such as setting up human-machine collaboration KPIs. At the policy level, "classified supervision" should be implemented, and AI ethical review such as content transparency labels should be strengthened for sensitive industries such as media, while a relaxed environment should be maintained for technology-intensive industries such as games and law to encourage pilot innovation. Multinational companies need to adjust their strategies based on regional characteristics - focus on technology ecosystem cooperation in the Japanese and Korean markets, and explore differentiated paths for large-scale applications in the Chinese market, such as vertical AI solutions. Research needs to further deconstruct the contribution of AIAD to segmented scenarios, such as customer service automation and supply chain forecasting, and develop a quantitative evaluation framework for HAICR to accurately identify the critical point of efficiency improvement in human-machine collaboration.

5.2. Consumer Trust Advice

Based on the analysis of consumer trust dynamics and AI tool selection, this article argues that governments should achieve dynamic regulatory adaptation and embed regional cultural assessment indicators, such as technology acceptance and privacy sensitivity, into

the regulatory framework. Strengthen compliance audits for policy-sensitive markets such as Germany, while focusing on ethical promotion rather than rigid regulation in culturally driven markets such as Japan.

At the same time, promote scenario-based trust building, add content tracing functions to image generation tools, and alleviate concerns about "deep fakes"; promote the disclosure of AI decision logic in the financial or educational fields, and replicate the successful model of the media industry to gain trust through high transparency.

5.3. Suggestions on Market Share

Based on the results of cluster analysis, this paper provides the following targeted development suggestions for different countries and industries. For countries with high AI adoption rates and high market shares, such as China and the United States, they should continue to increase their investment in the field of AI, especially in industries such as media, games, and manufacturing, to maintain their competitive advantage in the global market. These countries can further consolidate their leading position by strengthening talent training, promoting technological innovation, and formulating supportive policies. For countries with relatively low AI adoption rates, such as India and South Korea, it is recommended to strengthen infrastructure construction, improve education levels, and formulate incentive policies to promote the widespread application of AI technology in various industries, especially in areas with great potential such as law, healthcare, and education. These countries can learn from the successful experiences of other countries and formulate AI strategies suitable for their own development in combination with their own national conditions.

From an industry perspective, industries such as media, games, and retail show high AI adoption rates and market shares in multiple clusters. It is recommended that these industries continue to explore innovative applications of AI technology to improve user experience and operational efficiency. At the same time, for industries with low AI adoption rates, such as law and education, it is possible to strengthen AI training and knowledge popularization within the industry, lower the threshold for technology application, and promote the deep integration of AI technology with industry needs. In addition, the government and relevant institutions can provide customized policy support and financial assistance based on the characteristics of different industries to promote the balanced development of various industries. Through these measures, the level of global AI application can be further improved, and the digital transformation of various industries and the sustainable development of the economy can be promoted.

6. Conclusion

This study systematically reveals the multi-dimensional impact of AI technology on the economy, society and market, and verifies the transmission path from economy to society and then to market. This article can answer the three questions raised in the previous text. The results show that AIAD directly exacerbates job losses in highly automated industries such as manufacturing, but HAICR significantly alleviates this risk. Meanwhile, AIAD and HAICR have a positive effect on RI, but the effect depends on industry characteristics. For example, the legal and gaming industries have significant benefits. RS is

the core driving factor of CT, but its effect is affected by cultural background and tool transparency. Through cluster analysis, it is found that China and the United States hold dominant market positions in multiple industries such as media, games, and manufacturing, while India, South Korea, etc. need to strengthen infrastructure and policy support to catch up. The spatio-temporal dynamics are manifested as the cluster differences of different country-industry combinations. For example, Japan performs outstandingly in specific industries.

This study has the following limitations. First, it relies on a synthetic dataset based on publicly available reports and lacks authentic, fine-grained enterprise-level or individual-level data. Future research should seek to collaborate with industry partners to obtain real-world data for validation. Second, the model's predictive performance suggests the presence of unobserved confounding factors; future work could incorporate more complex models, such as Transformer-based time-series models, or a richer set of features. Finally, although the study covers a wide range of countries and industries, it does not yet encompass all regions and emerging sectors. Future interdisciplinary research should continue to expand the temporal scope and coverage of the data and explore methods for dynamic policy simulation.

References

1. Acemoglu, D.: Harms of ai. Working Paper No. w29247 (2021)
2. Acemoglu, D.: The simple macroeconomics of ai. *Economic Policy* 40(121), 13–58 (2025)
3. Bickley, S.J., Chan, H.F., Torgler, B.: Artificial intelligence in the field of economics. *Scientometrics* 127(4), 2055–2084 (2022)
4. Brynjolfsson, E., Rock, D., Syverson, C.: Artificial intelligence and the modern productivity paradox. *The Economics of Artificial Intelligence: An Agenda* pp. 23–57 (2019)
5. Calvano, E., et al.: Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review* 110(10), 3267–3297 (2020)
6. Cao, L.: Ai in finance: Challenges, techniques, and opportunities. *ACM Computing Surveys* 55(3), 1–38 (2022)
7. Chang, W.C.: A revised Girvan-Newman clustering algorithm for cooperative groups detection in programming learning. *Computer Science and Information Systems* 21, 491–505 (2024)
8. Czarnul, P., Matuszek, M.: Research on problem formulations in resource-aware problems across scientific domains and applications. *Computer Science and Information Systems* 22(2), 459–490 (2025)
9. Donika, A.D.: The ethics of artificial intelligence: A sociological approach. *Bioethics Journal* 16(2), 26–31 (2023)
10. Eisfeldt, A.L., Schubert, G., Zhang, M.B.: Generative ai and firm values. Working Paper No. w31222 (2023)
11. Federspiel, F., et al.: Threats by artificial intelligence to human health and human existence. *BMJ Global Health* 8(5), e010435 (2023)
12. Furman, J., Seamans, R.: Ai and the economy. *Innovation Policy and the Economy* 19(1), 161–191 (2019)
13. Gonzales, J.T.: Implications of ai innovation on economic growth: A panel data study. *Journal of Economic Structures* 12(1), 13–13 (2023)
14. Greenstein, S.: Preserving the rule of law in the era of artificial intelligence (ai). *Artificial Intelligence and Law* 30(3), 291–323 (2022)
15. Hampole, M., et al.: Artificial intelligence and the labor market. Working Paper No. w33509 (2025)

16. Hang, H., Chen, Z.: How to realize the full potentials of artificial intelligence (ai) in digital economy? a literature review. *Journal of Digital Economy* 1(3), 180–191 (2022)
17. Joyce, K., et al.: Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius* 7, 2378023121999581 (2021)
18. Leeuw, F.L.: The algorithmization of policy and society. *Artificial Intelligence and Evaluation* p. 242 (2025)
19. Makridakis, S.: The forthcoming artificial intelligence (ai) revolution: Its impact on society and firms. *Futures* 90, 46–60 (2017)
20. Makridis, C.A., Mishra, S.: Artificial intelligence as a service, economic growth, and well-being. *Journal of Service Research* 25(4), 505–520 (2022)
21. Masoumi, Z., van Genderen, J.: Artificial intelligence for sustainable development of smart cities and urban land-use management. *Geo-Spatial Information Science* 27(4), 1212–1236 (2024)
22. Milgrom, P.R., Tadelis, S.: How artificial intelligence and machine learning can impact market design. Working Paper No. c14008 (2018)
23. Parkes, D.C., Wellman, M.P.: Economic reasoning and artificial intelligence. *Science* 349(6245), 267–272 (2015)
24. Sartori, L., Theodorou, A.: A sociotechnical perspective for the future of ai: Narratives, inequalities, and human control. *Ethics and Information Technology* 24(1), 4–4 (2022)
25. Solos, W.K., Leonard, J.: On the impact of artificial intelligence on economy. *Science Insights* 41(1), 551–560 (2022)
26. Trammell, P., Korinek, A.: Economic growth under transformative ai. Technical Report No. w31815 (2023)
27. Warr, M.A., Rather, M.S.: Sociological analysis of artificial intelligence, benefits, concerns and its future implications. *International Journal of Indian Psychology* 12(3) (2024)
28. Whittlestone, J., Clarke, S.: Ai challenges for society and ethics. *The Oxford Handbook of AI Governance* (2022)
29. Xie, M.: Development of artificial intelligence and effects on financial system. *Journal of Physics: Conference Series* 1187(3) (2019)
30. Zhang, D., Lukovic, I.: Guest editorial: Explainable and trustworthy methods for next-generation artificial intelligence for reasonable data size. *Computer Science and Information Systems* 21(4), iii–iv (2024)

Na Tang is PhD student at the Department of Decision Sciences, Faculty of Business, University of Malaya and Lecturer of Big Data Management and Application at the School of Business, Geely University of China. Research areas include Machine learning, Big Data analytics, Algorithmic Consumer Behavior, Computational Marketing Science.

Yang Yang is Prof. PHD and Vice Dean at the School of Digital and New Media, Geely University of China. Research centers on brand management and computational communication, with current interest in digital brand management and data-intelligent communication.

Received: November 1, 2025; Accepted: June 25, 2026.

