Problem Chosen	2026HSB MCM/ICM Summary Sheet	Team Number
B		MI2601258

w This Is Title

Summary

Here is the summary of paper.

Please replace this text with your own summary. This summary should briefly describe the problem you are addressing, the methods you used, and the key results or conclusions you reached. Make sure to keep it concise and informative.

Keywords: KFC V50 Chicken

Contents

1 Intorduction

1.1 Background

In the contemporary era, artificial intelligence (AI) has emerged as one of the core domains of global technological competition, exerting profound and systemic influences on economic development, social progress, and national security. With the acceleration of a new wave of technological revolution and industrial transformation, AI technologies are fundamentally reshaping traditional industrial structures, modes of production, and governance systems, and have gradually become a key indicator of a nation's scientific strength and overall competitiveness.

Against this backdrop, countries around the world have elevated artificial intelligence to a strategic priority at the national level, continuously increasing investments in algorithmic research, computing infrastructure, data resource development, and the expansion of application scenarios, with the aim of securing a leading position in the global AI competitive landscape.

1.2 Problem Restatement

This study aims to quantitatively evaluate national artificial intelligence (AI) development capabilities, compare global competitiveness, and analyze future development trends through a systematic mathematical modeling framework. The problem is decomposed into four sequential and interrelated tasks:

Task 1: Factor Identification and Correlation Analysis

Relevant data are collected and integrated to identify the key factors influencing national AI development. These factors are quantified, and their intrinsic correlations and interaction mechanisms are analyzed using statistical and visualization methods.

Task 2: Comprehensive Evaluation and Ranking

Based on the quantified factors and their correlations obtained in Task 1, a multi-criteria evaluation model is constructed to assess and rank the AI competitiveness of ten selected countries.

Task 3: Competitiveness Trend Prediction

Using historical data from 2016 to 2025, the future evolution of AI development factors during the period 2026–2035 is predicted. The evaluation model established in Task 2 is then applied to analyze the dynamic changes in national competitiveness rankings over time.

Task 4: Optimal Fund Allocation Strategy

Under a fixed budget constraint of a 1 trillion yuan special fund, a multi-objective optimization model is developed to determine the optimal allocation of resources across AI development factors, with the goal of maximizing China's comprehensive AI competitiveness by 2035.

By sequentially accomplishing these tasks, this study provides a coherent framework for factor identification, comparative evaluation, future trend analysis, and strategic decision support in the global AI competition landscape.

1.3 Our Work

这是我们的工作介绍

2 Basic Assumption

To ensure the feasibility, consistency, and interpretability of the proposed models, the following basic assumptions are made.

► **Hypothesis 1: Assume that national AI development capability is a latent attribute that can be approximated by a finite set of observable and quantifiable indicators.**

Legitimacy: At the national level, AI development is manifested through measurable outcomes and resource inputs recorded in public statistics. Although the true capability cannot be observed directly, its major characteristics can be reasonably inferred from aggregated, quantifiable indicators.

► **Hypothesis 2: Assume that all indicators within the same evaluation year are cross-sectionally consistent.**

Legitimacy: Although data may be collected from slightly different release years, AI development is a long-term process. Minor temporal discrepancies do not significantly affect national-level competitiveness comparisons and help simplify the modeling process.

► **Hypothesis 3: Assume that the indicators are independent of each other in the weighting and evaluation stages.**

Justification : While interactions among indicators exist, explicitly modeling such dependencies would increase complexity and reduce interpretability. Treating indicators as independent avoids double counting and ensures the applicability of entropy-based and multi-criteria evaluation methods.

► **Hypothesis 4: Assume that the fundamental mechanisms of AI development remain stable during the forecasting and optimization period.**

Justification : National AI strategies, infrastructure construction, and talent cultivation generally evolve gradually. This stability makes trend-based prediction and investment optimization reasonable and analytically tractable.

3 Symbols

4 Data Explanation

4.1 Data Description and Sources

The dataset covers ten representative countries and consists of multiple quantitative indicators describing national AI development capability. For organizational clarity, the indicators were grouped into six dimensions reflecting different aspects of AI development. All data corresponded to the same evaluation period and were obtained from publicly available and widely recognized sources, ensuring cross-country comparability.

4.2 Data Preprocessing

Basic preprocessing was conducted prior to analysis. Minor missing or abnormal values were handled through reasonable estimation and consistency checks. All indicators were defined as benefit-type variables and were normalized to eliminate dimensional differences before being used in subsequent models.

5 Task 1: Identification and Structural Analysis of AI Development Factors

5.1 Factor System and Quantification

To quantitatively characterize national AI development capability, a multi-factor indicator system is constructed and organized into a standardized data matrix. Let

$$X = [x_{ij}] \in \mathbb{R}^{n\times p} \tag{1}$$

denote the normalized indicator matrix, where x_{ij} represents the standardized value of indicator j for country i, with $n = 10$ countries and $p = 24$ indicators.

To eliminate scale effects and ensure cross-country comparability, min–max normalization is applied:

$$x_{ij} = \frac{x_{ij}^{\text{raw}} - \min(x_j)}{\max(x_j) - \min(x_j)} \in [0, 1]. \tag{2}$$

The standardized matrix X serves as the common input for all subsequent structural analyses.

5.2 Correlation Structure among AI Development Factors

To explore linear associations among AI development factors, Pearson correlation coefficients are employed to quantify statistical dependence between indicators.

The correlation coefficient between factor j and factor k is defined as

$$r_{jk} = \frac{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2 \sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2}}, \tag{3}$$

forming the correlation matrix

$$R = [r_{jk}] \in \mathbb{R}^{p\times p}. \tag{4}$$

To emphasize statistically significant relationships, a strong-correlation edge set is defined as

$$\mathcal{E} = \{(j, k) \mid |r_{jk}| > \tau,\ j < k\}, \tag{5}$$

where τ denotes a predefined threshold.

The correlation results reveal a highly interconnected structure among AI development factors, with many indicator pairs exceeding the predefined threshold. This indicates pronounced co-movement rather than independent variation across dimensions.

As illustrated in the correlation heatmap, strong positive correlations form several contiguous blocks. Indicators related to government support, market scale, and investment intensity are closely linked with computing infrastructure and data-related factors, suggesting coordinated national development patterns. Talent-related indicators, including AI researchers, top AI scholars, and AI graduates, also exhibit strong mutual correlations and are closely associated with research output and enterprise activity.

By contrast, basic infrastructure penetration indicators display weaker correlations with advanced AI capability measures, implying that foundational infrastructure alone does not fully explain cross-country differences. Overall, the dense correlation structure highlights system-level coupling across multiple dimensions and suggests potential redundancy among indicators, motivating further structural analysis.

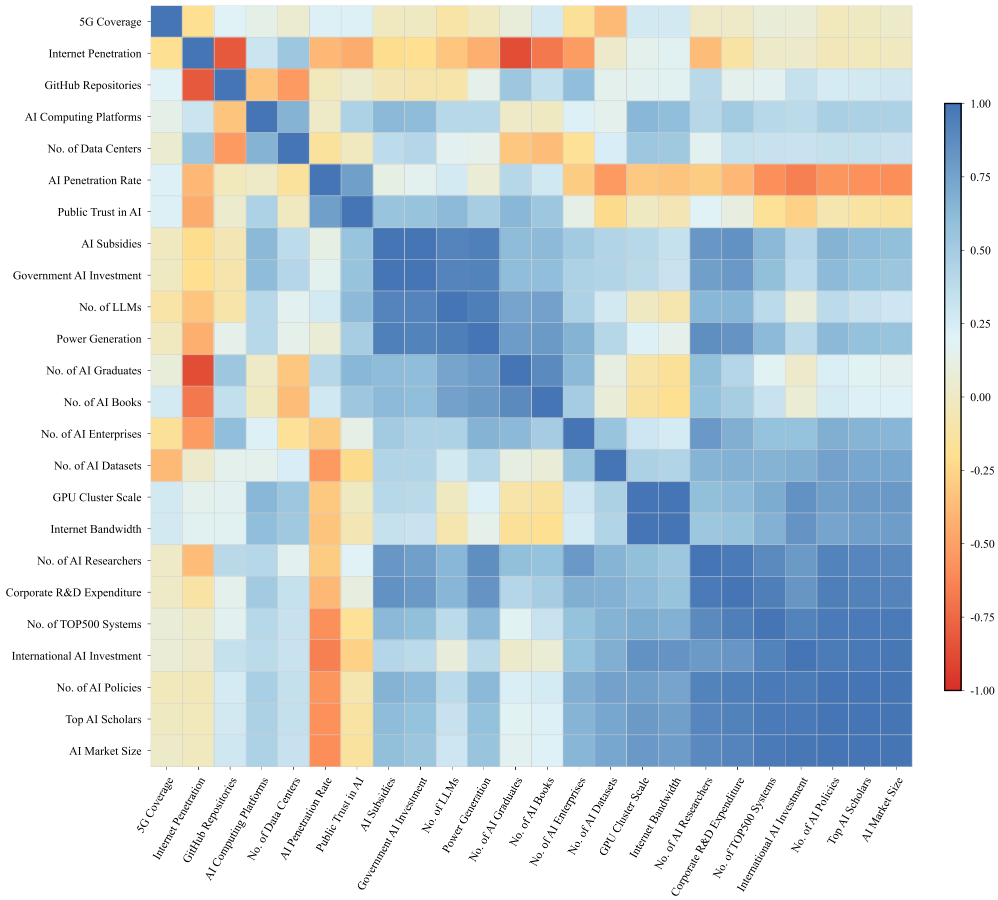

Figure 1: Correlation Heatmap of AI Development Factors

5.3 Structural Grouping via Hierarchical Clustering

To uncover higher-level structural organization beyond pairwise correlations, hierarchical clustering is conducted based on correlation-derived distances. The distance between factor j and factor k is defined as

$$d_{jk} = 1 - |r_{jk}|, \tag{6}$$

yielding the distance matrix $D = [d_{jk}]$.

Using the average linkage criterion, the distance between clusters C_a and C_b is computed as

$$D(C_a, C_b) = \frac{1}{|C_a||C_b|} \sum_{j \in C_a} \sum_{k \in C_b} d_{jk}. \tag{7}$$

The hierarchical clustering dendrogram reveals several coherent groups of AI development factors. Rather than being randomly aggregated, indicators are organized into structurally meaningful clusters reflecting consistent cross-country behavior.

One cluster primarily captures investment scale and market outcomes, including government and international AI investment, AI market size, and corporate R&D expenditure. Another cluster centers on computational capacity and infrastructure, grouping indicators such as GPU cluster scale, internet bandwidth, and data center availability. Talent- and knowledge-related indicators form a distinct cluster, reflecting strong alignment among human capital and research output factors.

In contrast, basic digital penetration indicators merge with other clusters at higher distance levels, indicating weaker direct association with advanced AI capability drivers. These results confirm that AI development factors form a multi-dimensional yet internally coordinated system.

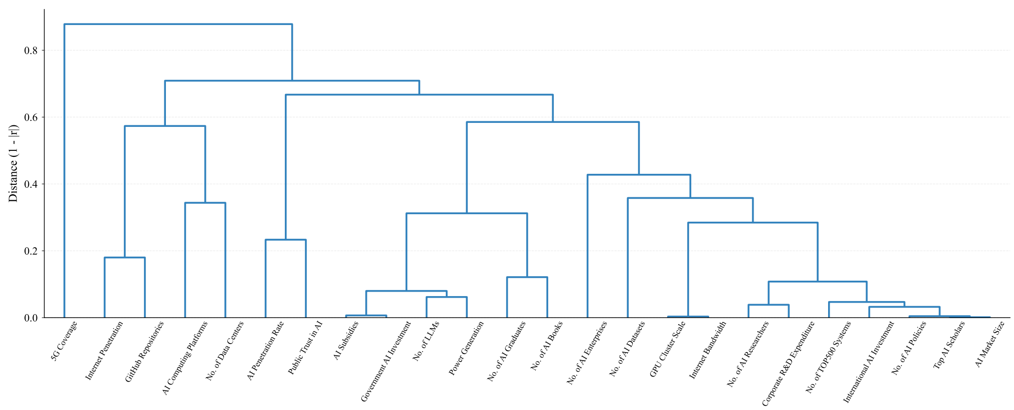

Figure 2: Hierarchical Clustering Dendrogram of AI Development Factors

5.4 Principal Component Analysis of Factor Structure

Given the high dimensionality of the indicator set, principal component analysis (PCA) is applied to extract a reduced set of comprehensive components while preserving dominant structural information.

The centered data matrix is

$$\tilde{X} = X - \mathbf{1}\bar{X}^T, \tag{8}$$

and the covariance matrix is

$$C = \frac{1}{n-1}\tilde{X}^T\tilde{X}. \tag{9}$$

Eigen-decomposition yields

$$C = V\Lambda V^T, \tag{10}$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$. The number of retained components m satisfies

$$\sum_{k=1}^{m} \frac{\lambda_k}{\sum_{j=1}^{p} \lambda_j} \geq \eta. \tag{11}$$

The PCA results show that variance is highly concentrated in the leading components. The first principal component captures a dominant share of total variance, indicating a strong common structural dimension across indicators. The cumulative variance exceeds 90% after four components, while additional components contribute only marginal explanatory power.

These findings demonstrate substantial redundancy in the original indicator set and justify representing the system with a low-dimensional structure. Retaining four principal components preserves essential information while significantly reducing dimensional complexity.

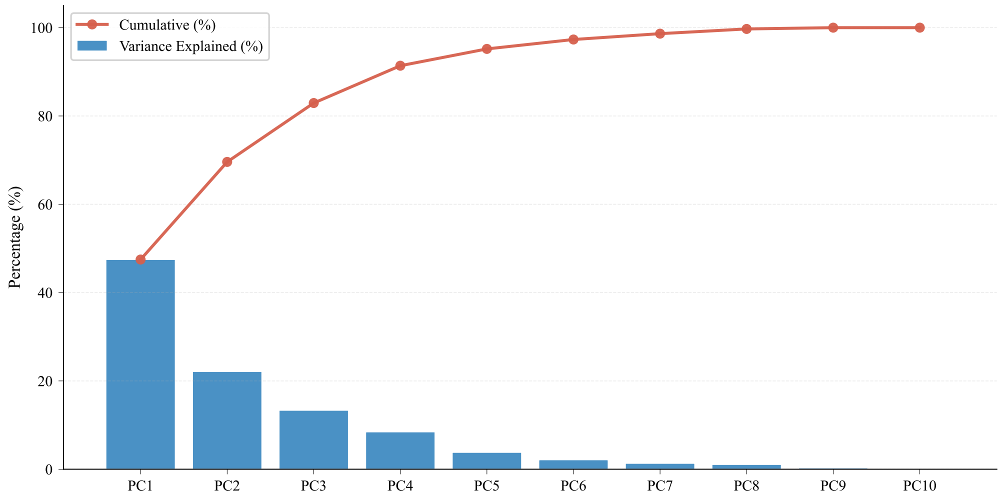

Figure 3: Variance Explained Plot for Principal Components

5.5 Relative Importance of AI Development Factors

Building on the reduced PCA representation, the relative importance of each indicator is quantified by integrating component loadings and variance contributions. The importance of factor j is defined as

$$I_j = \sum_{k=1}^{m} v_{jk}^2 \cdot \frac{\lambda_k}{\sum_{l=1}^{p} \lambda_l}. \tag{12}$$

The importance ranking indicates that explanatory power is concentrated in a limited subset of indicators. Factors related to human capital, research activity, investment intensity, and advanced infrastructure consistently exhibit higher importance scores.

High-importance indicators load strongly on the leading principal component, which dominates total variance. In contrast, lower-importance indicators primarily contribute to higher-order components and capture more localized variation. This concentration highlights the dominant drivers of cross-country AI capability differences and supports the use of weighted composite evaluation models.

5.6 Interaction Patterns among AI Development Factors

The strong-correlation network reveals a centralized and modular interaction structure among AI development factors. Indicators related to investment, computational capacity, and market scale occupy hub positions, linking multiple dimensions of the system.

Peripheral indicators, such as social perception and basic penetration measures, exhibit fewer strong connections, indicating weaker coupling with core AI capability drivers. Overall, the network structure confirms that AI development capability emerges from coordinated interactions among a limited set of dominant factors rather than isolated contributions.

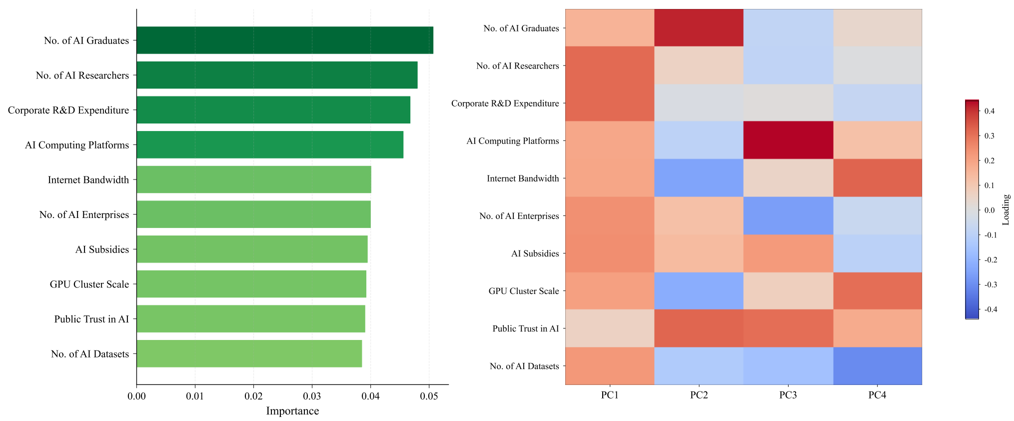

Figure 4: Factor Importance Ranking Bar Chart

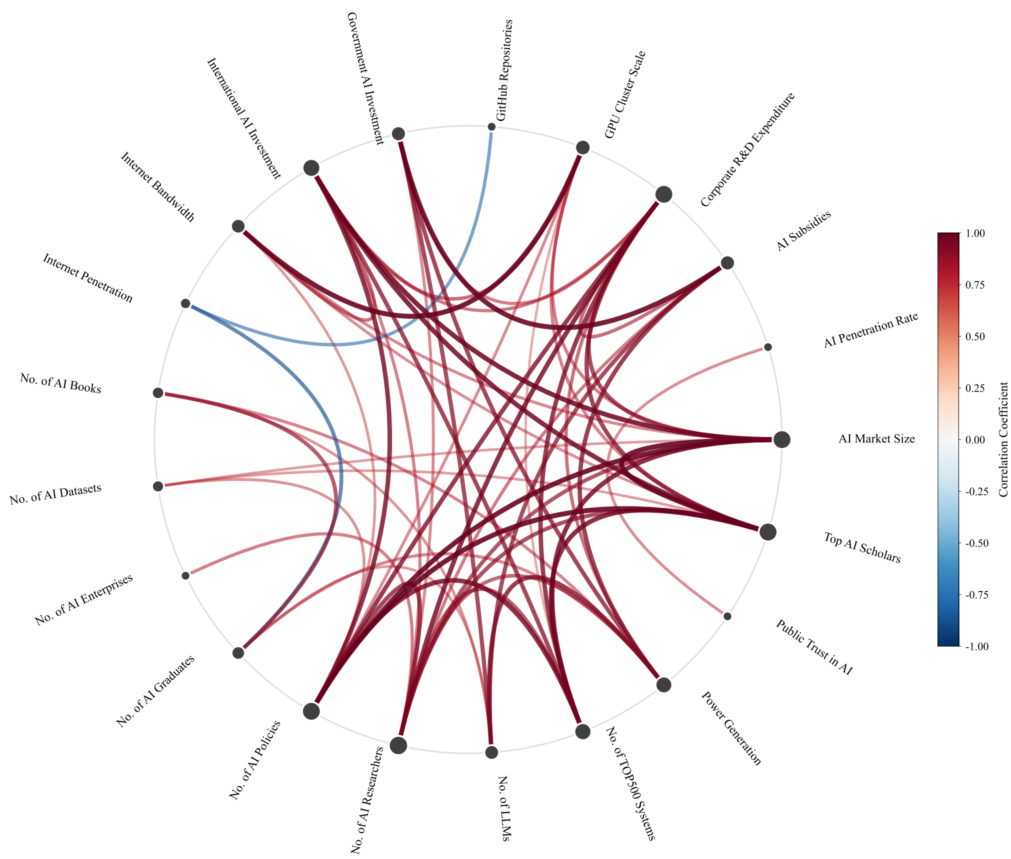

Figure 5: Strong-Correlation Network or Chord Diagram

5.7 Task 1 Summary

Task 1 provides a unified, data-driven structural analysis of national AI development capability by integrating correlation analysis, hierarchical clustering, PCA, and network analysis. The results demonstrate that AI development factors are highly interconnected and organized into coherent structural modules.

Dimensionality reduction reveals that a small number of dominant components capture most of the system's variance, while factor importance analysis identifies key indicators driving cross-country differences. Together, these findings establish a compact and consistent analytical foundation for the subsequent evaluation, ranking, and forecasting tasks.

6 Task 2: AI Development Capability Evaluation and 2025 Ranking

6.1 Model Overview

Based on the 24 indicators identified in Task 1, let the normalized indicator matrix be

$$X' = \left(x'_{ij}\right)_{n\times p}, \quad n = 10,\ p = 24, \tag{13}$$

where x'_{ij} denotes the normalized value of indicator j for country i. All indicators have been unified as benefit-type (larger values indicate stronger AI development capability) and normalized prior to this task.

The objective of Task 2 is to construct an *objective and reproducible* evaluation model to quantify national AI development capability and determine the 2025 competitiveness ranking of ten countries.

To reduce subjective bias and enhance robustness, an integrated evaluation framework combining the Entropy Weight Method (EWM), TOPSIS, and Grey Relational Analysis (GRA) is adopted.

6.2 Entropy-Based Weighting

The entropy weight method is derived from information theory and assigns indicator weights according to their dispersion across countries. Indicators with higher variability contain more effective information for discrimination and thus receive larger weights.

Define

$$p_{ij} = \frac{x'_{ij}}{\sum_{i=1}^{n} x'_{ij}}, \qquad 0\ln 0 := 0, \tag{14}$$

where p_{ij} represents the proportion of indicator j contributed by country i. The information entropy of indicator j is computed as

$$e_j = -k\sum_{i=1}^{n} p_{ij}\ln p_{ij}, \quad k = \frac{1}{\ln n}. \tag{15}$$

The corresponding entropy weight is given by

$$w_j = \frac{1-e_j}{\sum_{j=1}^{p}(1-e_j)}. \tag{16}$$

If $\sum_{i=1}^{n} x'_{ij} = 0$ for some indicator j, its weight is set to zero and the remaining weights are renormalized.

Result interpretation. The entropy-based weighting results indicate a clear concentration of information contribution among a limited subset of indicators. Specifically, frontier capability and high-level output indicators receive substantially larger weights than basic penetration metrics. The largest weights are assigned to *Large Models* ($w = 0.0946$), *International AI Investment* ($w = 0.0859$), *Top AI Scholars* ($w = 0.0699$), and *AI Market Size* ($w = 0.0698$), implying that cross-country differences are most strongly distinguished by cutting-edge innovation capacity, global resource linkage, and market maturity. In contrast, *Internet Penetration* ($w = 0.0063$), *AI Penetration* ($w = 0.0084$), and *5G Coverage* ($w = 0.0087$) carry relatively small weights, suggesting that foundational digital access has partially converged among major economies and is less decisive in differentiating AI competitiveness at the current stage.

Overall, the weight structure is consistent with the structural findings in Task 1: a small set of dominant factors explain most of the cross-country variance, while low-weight indicators provide complementary but limited marginal discrimination.

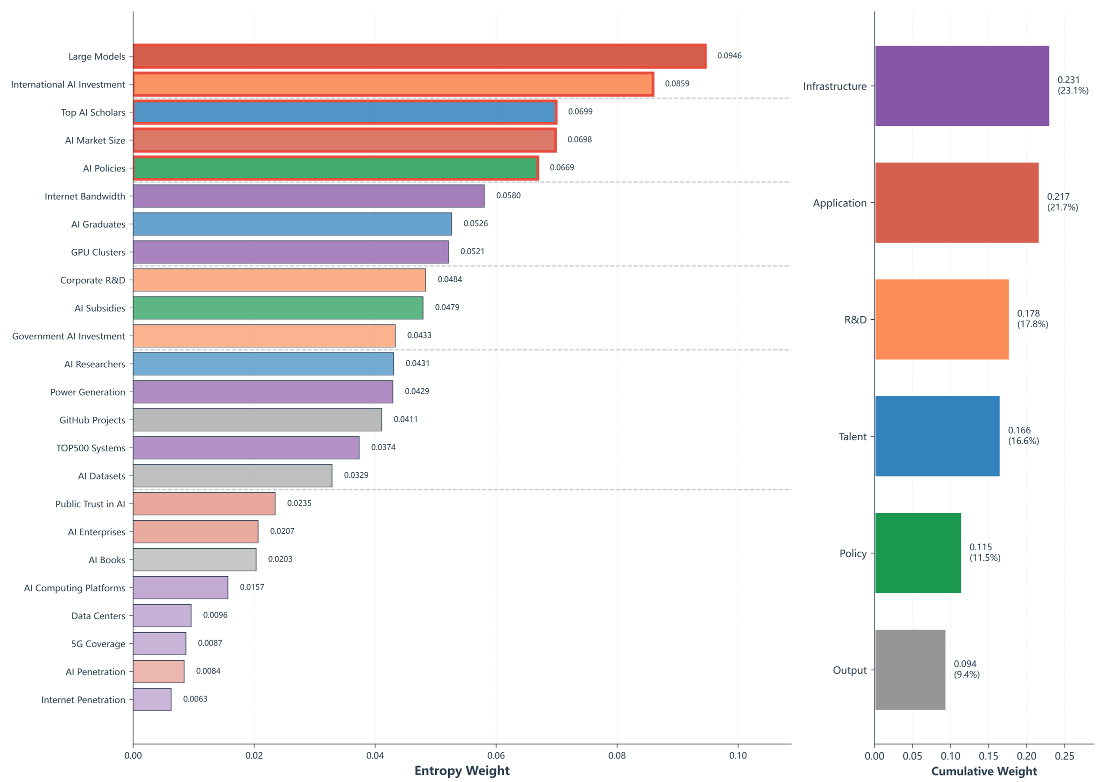

Figure 6: Indicator Weights Distribution (Entropy Weight Method)

6.3 TOPSIS-Based Comprehensive Evaluation

TOPSIS aggregates weighted indicators into a single competitiveness score by comparing each country with an ideal reference. The method assumes that the optimal country should be closest to the

Table 1: Entropy Weights of Representative Indicators (Top and Bottom)

Indicator	Entropy e_j	Redundancy $(1 - e_j)$	Weight w_j
Large Models	0.2144	0.7856	0.0946
International AI Investment	0.2862	0.7138	0.0859
Top AI Scholars	0.4192	0.5808	0.0699
AI Market Size	0.4201	0.5799	0.0698
AI Policies	0.4444	0.5556	0.0669
AI Penetration	0.9300	0.0700	0.0084
5G Coverage	0.9275	0.0725	0.0087
Internet Penetration	0.9478	0.0522	0.0063

positive ideal solution and farthest from the negative ideal solution.

The weighted decision matrix is defined as

$$v_{ij} = w_j x'_{ij}. \tag{17}$$

For each indicator j, the positive and negative ideal components are defined as

$$A_j^+ = \max_i v_{ij}, \quad A_j^- = \min_i v_{ij}, \qquad j = 1, \ldots, p, \tag{18}$$

yielding the ideal vectors $A^+ = (A_1^+, \ldots, A_p^+)$ and $A^- = (A_1^-, \ldots, A_p^-)$.

The distances from country i to the ideal solutions are computed as

$$D_i^\pm = \sqrt{\sum_{j=1}^{p} (v_{ij} - A_j^\pm)^2}, \tag{19}$$

and the comprehensive competitiveness score is defined as

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-}. \tag{20}$$

Result interpretation. The entropy-weighted TOPSIS results demonstrate strong heterogeneity in national AI development capability in 2025. The United States ranks first with $C = 0.6407$, followed by China with $C = 0.5104$, forming a clear leading tier. India ($C = 0.2098$) and the UAE ($C = 0.1836$ constitute the next tier, while the United Kingdom, South Korea, France, Japan, and Germany cluster in a compact mid-lower band ($C \approx 0.062$–0.100). Canada ranks last with $C = 0.0414$.

This "head–middle–tail" pattern indicates that top countries outperform others simultaneously across multiple high-weight indicators, while the separation among mid-ranked countries is relatively small and therefore sensitive to structural emphasis across dimensions.

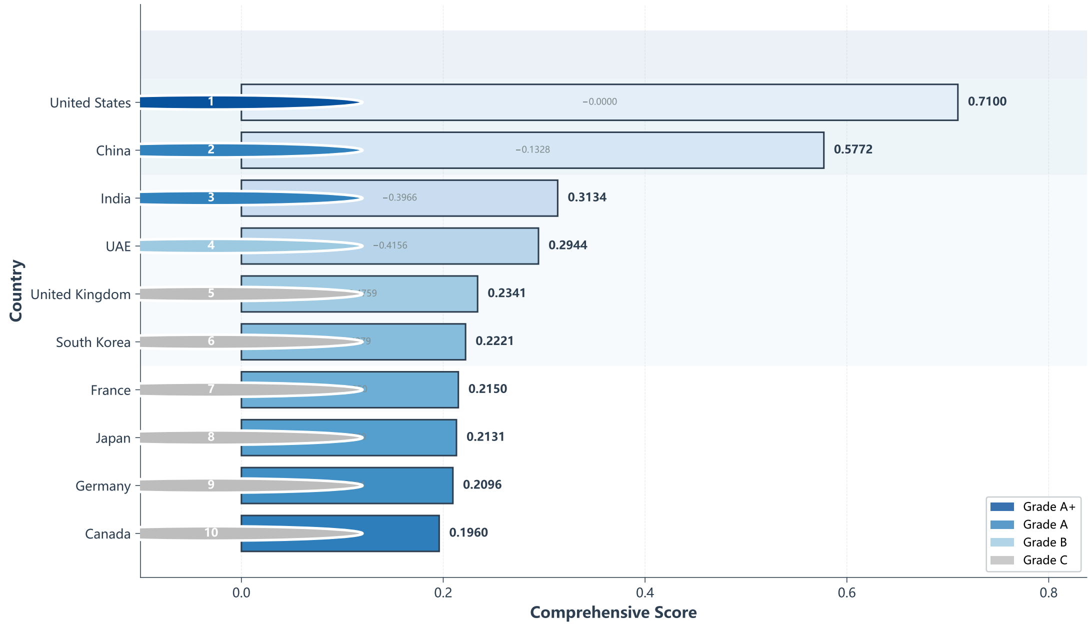

Figure 7: AI Competitiveness Ranking in 2025 (TOPSIS Score)

Table 2: TOPSIS Comprehensive Evaluation Results (2025)

Country	D_i^+	D_i^-	TOPSIS Score C_i	Rank
United States	0.0978	0.1744	0.6407	1
China	0.1320	0.1377	0.5104	2
India	0.2006	0.0533	0.2098	3
UAE	0.1998	0.0449	0.1836	4
United Kingdom	0.2040	0.0225	0.0995	5
South Korea	0.2063	0.0165	0.0740	6
France	0.2065	0.0153	0.0688	7
Japan	0.2064	0.0141	0.0640	8
Germany	0.2063	0.0138	0.0625	9
Canada	0.2091	0.0090	0.0414	10

6.4 Structural Validation by Grey Relational Analysis

Grey Relational Analysis is suitable for small-sample and partially known systems. It evaluates the similarity of development patterns by comparing the geometric proximity of indicator sequences.

Let the reference sequence be the ideal profile

$$v_{0j} = \max_i v_{ij}, \qquad j = 1, \dots, p. \tag{21}$$

Define the absolute deviation

$$\Delta_{ij} = |v_{0j} - v_{ij}|, \tag{22}$$

and let

$$\Delta_{\min} = \min_{i,j} \Delta_{ij}, \qquad \Delta_{\max} = \max_{i,j} \Delta_{ij}. \tag{23}$$

The grey relational coefficient is defined as

$$\xi_{ij} = \frac{\Delta_{\min} + \rho\Delta_{\max}}{\Delta_{ij} + \rho\Delta_{\max}}, \quad \rho = 0.5, \tag{24}$$

and the grey relational degree of country i is calculated as

$$\gamma_i = \frac{1}{p}\sum_{j=1}^{p} \xi_{ij}. \tag{25}$$

Result interpretation. GRA is employed as an independent structural validation of the TOPSIS-based ranking. The resulting grey relational degrees preserve the same leading pair as TOPSIS: the United States ranks first ($\gamma = 0.7793$) and China ranks second ($\gamma = 0.6440$). India ($\gamma = 0.4170$) and the UAE ($\gamma = 0.4052$) remain in the upper-middle group. Within the mid-ranked countries, slight differences emerge (e.g., South Korea ranks above the United Kingdom in GRA), indicating that some countries may have more coherent indicator structures even when their absolute TOPSIS scores are close.

Overall, the close alignment between TOPSIS and GRA supports that the ranking is not an artifact of a single evaluation logic, but reflects stable and systematic capability differences across countries.

Table 3: Grey Relational Degrees and Rankings

Country	Grey Relational Degree γ_i	Rank
United States	0.7793	1
China	0.6440	2
India	0.4170	3
UAE	0.4052	4
South Korea	0.3703	5
United Kingdom	0.3686	6
Japan	0.3623	7
France	0.3612	8
Germany	0.3567	9
Canada	0.3506	10

6.5 Fusion Ranking and Reliability Framework

To integrate distance-based performance and structural similarity, an equal-weight fusion score is defined as

$$S_i = \frac{C_i + \gamma_i}{2}. \tag{26}$$

The fusion score S_i is used to determine the final AI competitiveness ranking for 2025.

Ranking consistency is quantified using Spearman's rank correlation coefficient

$$\rho_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \tag{27}$$

where d_i denotes the rank difference of country i between two ranking methods. Robustness is further assessed by perturbing a single indicator weight w_j by $\pm\alpha$ (e.g., $\alpha = 30\%$), followed by weight renormalization and re-ranking.

Final ranking and reliability analysis. The equal-weight fusion of TOPSIS and GRA produces the final 2025 competitiveness ranking. The top two countries (United States and China) remain unchanged, and most mid-ranked countries show only minor rank adjustments, indicating a stable ordering. Method consistency is extremely high: the rank correlation between TOPSIS scores and grey relational degrees achieves Spearman $\rho_s = 0.9758$ with statistical significance ($p < 0.001$), confirming that the evaluation outcome is robust to the choice of ranking mechanism. Sensitivity results further indicate limited rank volatility under weight perturbations: most countries exhibit rank ranges of 0–1, and only a few countries show moderate variation (e.g., France with range 3), suggesting overall model stability.

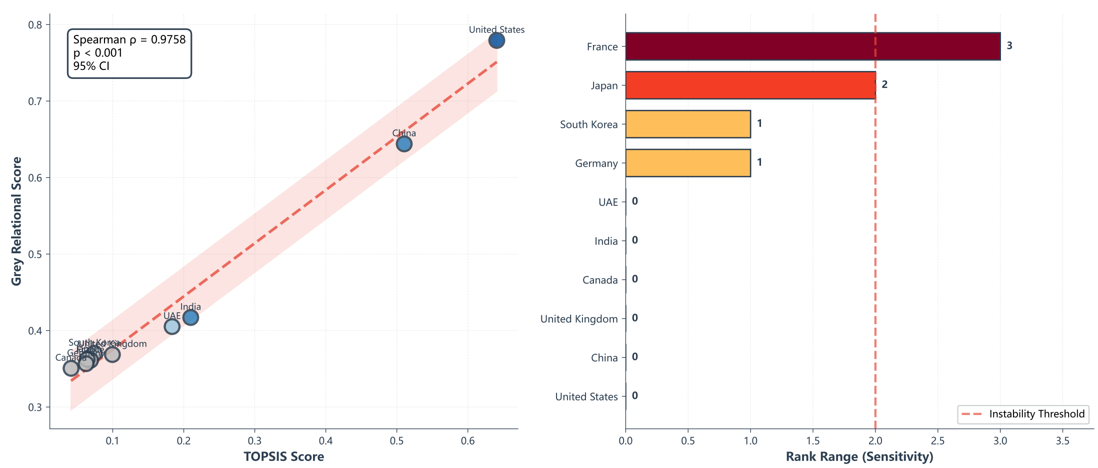

Figure 8: Validation and Robustness Analysis (Method Consistency and Sensitivity)

Structural explanation by six dimensions (optional). To interpret why top-ranked countries outperform others, the six-dimensional scores provide a compact decomposition. The United States exhibits strong advantages in *R&D* (0.8307) and *Infrastructure* (0.8714), while China shows prominent strengths in *Application* (0.7344) and *Policy* (0.7980). India displays relatively high *Output* (0.5889) and moderate *Application* (0.4011), whereas the UAE relies more on *Infrastructure* (0.6453) and policy-related support. These complementary structures help explain why the leading tier remains stable while mid-ranked countries cluster closely.

Table 4: Final AI Competitiveness Ranking (2025) by Fusion Score

Country	TOPSIS Rank	GRA Rank	TOPSIS C_i	GRA γ_i	Fusion Score S_i	Grade
United States	1	1	0.6407	0.7793	0.7100	A+
China	2	2	0.5104	0.6440	0.5772	A
India	3	3	0.2098	0.4170	0.3134	A
UAE	4	4	0.1836	0.4052	0.2944	B
United Kingdom	5	6	0.0995	0.3686	0.2341	C
South Korea	6	5	0.0740	0.3703	0.2221	C
France	7	8	0.0688	0.3612	0.2150	C
Japan	8	7	0.0640	0.3623	0.2131	C
Germany	9	9	0.0625	0.3567	0.2096	C
Canada	10	10	0.0414	0.3506	0.1960	C

Table 5: Ranking Sensitivity under Weight Perturbation (Rank Range)

Country	Rank Range
United States	0
China	0
India	0
UAE	0
United Kingdom	0
Canada	0
Germany	1
South Korea	1
Japan	2
France	3

6.6 Overall Summary of Task 2

Task 2 establishes a comprehensive evaluation framework for national AI development capability by integrating objective weighting, distance-based evaluation, and structural validation. Using entropy-based weights, the TOPSIS model yields a clear 2025 competitiveness hierarchy, with the United States and China forming the leading tier, followed by India and the UAE, and a compact cluster of mid-ranked countries.

The GRA results provide an independent validation of the TOPSIS ordering, and the extremely high rank correlation (Spearman $\rho_s = 0.9758$, $p < 0.001$) confirms strong methodological consistency. Sensitivity analysis further indicates that rankings remain stable under reasonable perturbations of indicator weights. Overall, the final fusion ranking offers a transparent and reproducible benchmark for 2025 AI competitiveness, serving as the baseline for forecasting and policy analysis in subsequent tasks.

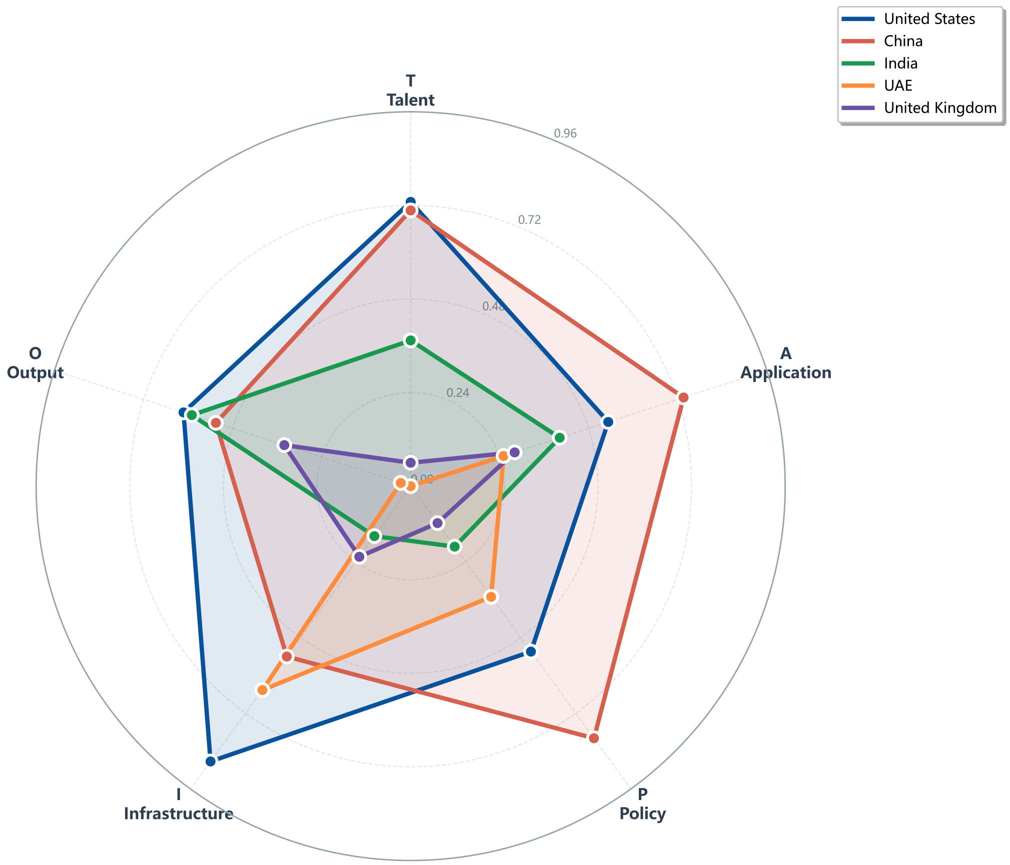

Figure 9: Six-Dimensional Capability Profiles of Representative Countries

7 Task 3: AI Competitiveness Ranking Forecast (2026–2035)

7.1 Model Objective and Overall Framework

Building upon the indicator system and structural insights identified in Task 1, as well as the objective evaluation framework established in Task 2, Task 3 focused on forecasting the evolution of national AI competitiveness rankings over the period 2026–2035. The primary objective was to construct a *forecast–evaluation integrated pipeline* that remained fully reproducible, interpretable, and strictly consistent with the previously defined evaluation model.

Let $x_{i,j,t}$ denote the observed value of indicator j for country i in year t. Using the complete panel data from 2016 to 2025, Task 3 was implemented through three sequential stages:

1. indicator-level time series forecasting for each country and indicator;
2. annual comprehensive evaluation using the fixed weighting and TOPSIS framework from Task 2;
3. ranking evolution and driver analysis across the forecast horizon.

This design ensured that all ranking changes were driven exclusively by data-inferred temporal dynamics, without introducing any exogenous assumptions or subjective adjustments.

7.2 Data Structure and Notation

The data used in Task 3 followed a balanced panel structure:

- Countries: $i = 1, \ldots, n$, with $n = 10$;
- Indicators: $j = 1, \ldots, p$, with $p = 24$;
- Historical years: $t \in \{2016, \ldots, 2025\}$;
- Forecast years: $t \in \{2026, \ldots, 2035\}$.

For each forecast year t, the predicted indicator matrix was denoted as

$$\hat{X}_t = \left(\hat{x}_{i,j,t}\right)_{n\times p}, \tag{28}$$

which served as the direct input to the annual evaluation procedure.

To preserve physical interpretability and comparability across indicators, indicator-specific constraints were imposed:

- non-negativity constraints for scale and quantity indicators;
- bounded intervals for ratio- or proportion-type indicators.

7.3 Indicator Forecasting Methodology

7.3.1 Baseline Grey Forecasting Model

Given the limited historical sample size available for each indicator sequence, the Grey Model GM(1,1) was adopted as the baseline forecasting approach. For a fixed country–indicator pair (i, j), the original sequence was defined as

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(T)\right), \quad T = 10. \tag{29}$$

After applying first-order accumulated generation (AGO), a first-order differential equation was fitted and solved to obtain the corresponding time response function. The predicted original sequence was then recovered via inverse AGO. The GM(1,1) model was selected for its robustness in small-sample settings and its suitability for capturing monotonic development trends.

7.3.2 Engineering Constraints and Stabilization

To ensure numerical stability and realistic forecasts across heterogeneous indicators, several unified treatments were applied:

- non-negative translation for near-zero or zero-valued series prior to model fitting;
- boundary truncation of forecast values according to indicator type;
- optional logarithmic transformation for indicators with large magnitude dispersion.

All transformations were rule-based and reversible, and no country-specific adjustments were introduced at any stage.

7.3.3 Backtesting and Fallback Strategy

Model suitability was evaluated via one-step-ahead backtesting. Specifically, data from 2016–2024 were used for training, and the year 2025 was forecast for validation. Prediction error was measured using the mean absolute percentage error (MAPE).

Result placeholder.

To be added: summary statistics of backtesting errors across all country–indicator pairs, using data from `forecast_diagnostics.csv`.

Suggested table: Three-line table reporting the distribution of MAPE values and the proportion of series exceeding the predefined threshold.

When the GM(1,1) model failed to satisfy accuracy or stability requirements, a fallback predictor (linear trend or Holt exponential smoothing) was automatically activated, subject to the same boundary constraints.

7.4 Annual Evaluation Using Fixed Weights and TOPSIS

7.4.1 Weight Inheritance Principle

Let $W = (w_1, \ldots, w_p)$ denote the entropy-based indicator weights obtained in Task 2. Throughout Task 3, these weights were held constant across all forecast years to ensure temporal comparability and methodological consistency.

7.4.2 Yearly TOPSIS Evaluation

For each forecast year t, the predicted matrix $\hat{X}_t$ was normalized and weighted using the same TOPSIS procedure as in Task 2. The relative closeness score for country i in year t was denoted as

$$C_{i,t} \in [0, 1]. \tag{30}$$

Result placeholder.

To be added: annual TOPSIS scores for all countries during 2026–2035, based on `topsis_scores_2026_2035.cs`

Suggested table: Three-line table showing selected years (e.g., 2026, 2030, 2035) with corresponding scores for each country.

Suggested figure: Line plot illustrating the temporal evolution of $C_{i,t}$ for major economies.

7.5 Ranking Generation and Evolution Analysis

For each year t, countries were ranked in descending order of $C_{i,t}$:

$$\text{Rank}_{i,t} = \text{rankdesc}\big(C_{:,t}\big)\,. \tag{31}$$

Result placeholder.

To be added: complete ranking results for 2026–2035 using `rankings_2026_2035.csv`.

Suggested table: Three-line table summarizing average rank, maximum rise/fall, and rank standard deviation for each country.

Suggested figure: Ranking evolution plot or heatmap based on `rankings_2026_2035_wide.csv`.

7.6 Driving Factor Analysis of Ranking Changes

To interpret ranking dynamics, changes in the TOPSIS score were decomposed with respect to high-weight indicators. For a fixed country, score variations were analyzed in conjunction with changes in weighted normalized indicator values.

Result placeholder.

To be added: identification of key indicators contributing to score changes for representative countries.
Data source: weighted normalized matrices derived from predicted indicators and Task 2 weights.
Suggested figure: Contribution bar charts or radar plots highlighting dominant drivers.

7.7 Model Validation and Robustness Checks

Model robustness was examined from both the forecasting and evaluation perspectives:

- comparison between predicted and observed 2025 evaluation scores to validate the end-to-end pipeline;
- sensitivity of rankings to fallback thresholds and minor perturbations in high-weight indicators.

Result placeholder.

To be added: robustness assessment results using sensitivity scenarios.
Suggested table: Three-line table reporting rank correlations under different perturbation settings.

7.8 Summary of Task 3

Task 3 extended the static evaluation framework into a dynamic forecasting context by integrating indicator-level time series prediction with a fixed and validated evaluation model. The resulting rankings reflected endogenous development trajectories inferred from historical data, rather than externally imposed assumptions.

Overall, the proposed pipeline preserved methodological continuity across tasks, maintained interpretability at both the indicator and system levels, and provided a quantitative basis for analyzing future shifts in global AI competitiveness. These results established a consistent foundation for the subsequent policy-oriented optimization analysis.

8 Task 4