**S&P 500 independent component analysis flow chart**

|  |
| --- |
| **S& P 500 Stock Dataset** Historical price data | Volume data |

↓

|  |
| --- |
| Data Preprocessing Yield calculation + standardization |

↓

|  |  |  |
| --- | --- | --- |
| **PCA** | **VS** | **ICA** |

↓

|  |
| --- |
| **Step 4: Risk factor identification** Four independent risk factors were accurately identified IC1: Peak risk factor  IC2: Volatility risk factor IC3: Distribution morphological factor  IC4: Yield risk factor |

↓

|  |
| --- |
| **Step5: K-Means Clustering**  Building Five portfolio clusters Cluster 0: Low risk and robust Cluster 1: High growth technology  Cluster 2: Period value type Cluster 3: Defensive revenue type  Group 4: Resource sensitive |

↙ ↘

|  |  |  |
| --- | --- | --- |
| Statistical significance test ANOVA test Quantitative comparison in 7 dimensions ICA vs PCA: 0.231 vs 0.150 |  | Portfolio backtest verification Risk control strategy Quantitative transaction verification Energy and technology sector, differentiated risk characteristics |

↘ ↙

|  |
| --- |
| Core achievements and applications Portfolio management optimization | Quantitative identification of risk factors Providing Support for Scientific Decision-Making |