**北 京 师 范 大 学**

**硕士研究生学位论文开题报告**

**Beijing Normal University**

**Graduate Student**

**Thesis Proposal Report**

姓 名Name SHURJO FAZLE MEHRAB

学 号Student I.D. 202229030001

导 师Supervisor 郑飞虎

系 所School Business School

专 业 Major World Economy

研究方向 Research Field World Economy and China

入学时间Entry Date 2023 年 10 月 28

**论文题目Thesis Title:**

**Research on asset specificity and firm's operations: classification of machine learning model**

1. **Theoretical Basis立论依据**

|  |
| --- |
| **(Research Significance, Domestic and International Research Analysis选题的研究意义、国内外研究现状分析)**  Determining whether an asset or resource is specific or non-specific to a firm's operations" refers to the classification or categorization of assets based on their degree of specialization or adaptability within a particular company's operations. This classification aims to assess how uniquely tailored or general-purpose an asset is in fulfilling the specific needs of a firm.  **Specific Asset**: An asset is considered specific if it is uniquely suited, tailored, or designed to meet the specific requirements of a particular firm's operations. These assets may have attributes, characteristics, or functionalities that are highly specialized or custom-made for the firm's processes, products, or services. Specific assets might include specialized machinery, proprietary technologies, or unique intellectual property directly aligned with the firm's operations.  **Non-Specific Asset**: Conversely, a non-specific asset refers to resources that are more generic, versatile, or widely applicable across various firms or industries. These assets are not specifically tailored or customized for the unique needs of a particular firm and may have more general-purpose applications. Non-specific assets could include standard office equipment, common software applications, or widely available raw materials.  The determination of asset specificity is crucial for firms in decision-making processes, resource allocation, strategic planning, and assessing competitive advantages. Understanding which assets are specific or non-specific helps firms identify their core competencies, evaluate their unique advantages, and make informed decisions regarding investment, outsourcing, or leveraging their distinct assets for competitive positioning.  If we consider a real-life scenario in the context of a manufacturing company that produces specialized machinery used in the semiconductor industry. The company aims to determine whether a particular piece of machinery is specific or non-specific based on various contract, buyer-seller, technical, and contractual information available in their dataset. In the context of semiconductor fabrication equipment, machinery that could be considered non-specific or generic might include:   1. **Basic Manufacturing Tools:**    * Tools such as standard cutting machines, basic drilling equipment, or common grinding tools might be considered non-specific as they are widely used across various manufacturing industries and may not be specifically tailored for semiconductor fabrication. 2. **Generic Testing Equipment:**    * Some testing equipment, like basic multimeters or oscilloscopes commonly found in electronics manufacturing, might not be specifically designed for semiconductor fabrication processes and could be considered non-specific. 3. **Common Assembly Line Conveyors:**    * Standard conveyor systems used in manufacturing assembly lines that are not specifically designed or optimized for semiconductor production might fall into the non-specific category.   In a semiconductor manufacturing firm, specialized machinery could be:   1. **Photolithography Systems:**    * These machines are essential for printing intricate circuit patterns onto semiconductor wafers using photomasks. They utilize light exposure to transfer the pattern onto the wafer's surface, a critical step in the semiconductor fabrication process. 2. **Etching Equipment:**    * Specialized etching machines are used to remove material from the semiconductor wafer selectively. They enable precise etching of the semiconductor material to create circuits and patterns according to the design specifications. 3. **Deposition Tools (Chemical Vapor Deposition/CVD, Physical Vapor Deposition/PVD):**    * Deposition tools are used to deposit thin layers of materials onto the semiconductor wafer's surface. Chemical and physical vapor deposition methods help in creating uniform layers of materials like silicon, metals, or insulators. |

**二、Research Methods研究方案**

|  |
| --- |
| Research Methods: The research methods include data collection, data preprocessing, machine learning model selection, and evaluation. Python will be employed as the primary programming language, and libraries such as Scikit-Learn and Pandas will be used for data analysis. The choice of models and evaluation metrics will depend on the research objectives.  Technology Roadmap: The technology roadmap encompasses the use of machine learning algorithms to analyze the dataset. Computational resources and tools such as Jupyter Notebooks and cloud-based platforms will be considered. Appropriate libraries for data visualization will also be utilized.  Feasibility Analysis: While the research is technically feasible, several challenges may arise, including data availability and data privacy concerns. Ethical considerations and potential limitations will be addressed. A detailed feasibility analysis will be undertaken to ensure the smooth progression of the research.    **Steps in Model Building:**   1. **Data Preprocessing:** This involves cleaning the dataset, handling missing values, encoding categorical variables, and splitting data into training and testing sets. 2. **Feature Selection:** Identifying the most relevant features (independent variables) that strongly influence asset specificity using techniques like feature importance analysis or domain knowledge. 3. **Model Training:** Training each of these algorithms using the training dataset, fitting them to learn patterns and relationships between features and asset specificity. 4. **Model Evaluation:** Evaluating the performance of each model using the test dataset, assessing metrics like accuracy, precision, recall, and F1-score to determine which algorithm performs best in predicting asset specificity. 5. **Model Comparison and Selection:** Comparing the performance of Logistic Regression, Decision Trees, Random Forest, and SVM to choose the most suitable model for predicting asset specificity based on our dataset. Each of these model are explained in details below: 6. **Logistic Regression:**    * **What it does:** Logistic Regression is a type of algorithm used for binary classification tasks, where the goal is to predict outcomes that have two possible values (e.g., yes/no, true/false).    * **How it works:** It models the relationship between the independent variables and the probability of a specific outcome using a logistic function. It estimates the probability of an event occurring by fitting a curve to the data.    * **Example:** Predicting whether a student will pass (yes) or fail (no) in an exam based on study hours, past grades, and attendance. 7. **Decision Trees:**    * **What they do:** Decision Trees are a tree-like structure used for both classification and regression tasks.    * **How they work:** They make decisions by splitting the dataset into smaller subsets based on different attributes. Each node in the tree represents a decision based on a feature, leading to a final decision or prediction at the tree's leaves.    * **Example:** Classifying whether a fruit is an apple or an orange based on features like color, size, and shape. 8. **Random Forest:**    * **What it is:** Random Forest is an ensemble learning method that uses multiple Decision Trees for classification or regression.    * **How it works:** It creates a 'forest' of decision trees by training several trees on random subsets of the dataset and combining their predictions to make a final prediction. It reduces overfitting and increases accuracy compared to a single Decision Tree.    * **Example:** Predicting whether an email is spam or not by considering various features like subject, sender, and content. 9. **Support Vector Machines (SVM):**    * **What they do:** SVMs are used for classification tasks, especially in cases where there is a clear separation between classes.    * **How they work:** SVMs find a hyperplane (boundary) that best separates different classes in the dataset. It maximizes the margin between the classes, aiming to create the widest possible gap.    * **Example:** Classifying whether a patient has a particular disease or not based on medical test results, age, and other health indicators.   To build an asset specificity classification machine learning (ML) model, it's essential to select relevant features (independent variables) and define the target variable (dependent variable) for classification. From the provided dataset, I'll identify potential independent and dependent variables:  **Dependent Variable (Target):**   * **Asset Specificity Classification:** This variable will indicate whether an asset is specific or non-specific based on the provided features.   **Independent Variables (Features):** Considering the context of asset specificity, I'll select relevant features that might influence asset classification:   1. **Contract Details:**    * Contract Number    * Delivery Date    * Amount    * Year    * Classification of the Contract    * Contract Type A    * Contract Type B 2. **Buyer-Seller Information:**    * Name of the Buyer    * Location of the Buyer    * Classification of the Buyer    * Name of the Seller    * Seller Category A    * Seller Category B    * Seller Category C    * Seller Ownership 3. **Technical and Outcome Details:**    * Technical Field A    * Technical Field B    * Outcome: Patent    * Outcome: Know-how    * Outcome with No IP    * Knowstyle    * Intellectual Property A    * Intellectual Property B 4. **Project Information:**    * Name of the Project    * If the Project is an R&D Mission    * If the Project is a Technical Consultation    * If the Project is an IP Exchange 5. **Additional Relevant Features:**    * Export    * Export to OECD    * Buyer's Country    * If the Buyer is a Chinese Organization    * Socioeconomic Goals    * Payment Method   **Excluded Variables:**   * Variables like Date of Sitting, Time, Seller Registration Number, Seller Organization Number, etc., might not directly contribute to the classification of asset specificity and are excluded from the selected features. |
| **3．Distinguished Features and Innovation 本研究的特色与创新之处**   * The incorporation of machine learning techniques into the research sets it apart. This innovative approach expands the analytical horizons by leveraging the power of data-driven insights. By applying machine learning classification algorithms, the study will explore patterns and correlations within the context of asset specificity and open innovation. This technological integration marks a significant step forward in the field.   **4．Expected Progress and Results 预期的论文进展和成果**  The initial stages of this research will encompass an extensive exploration of the available data and a thorough process of data visualization. These preliminary activities are planned to unfold over the first few weeks of the project. The primary objective during this phase is to gain a deep and nuanced understanding of the dataset at hand, which forms the foundation for all subsequent analyses and machine learning endeavors.  The data exploration phase involves comprehensive data cleansing, validation, and organization. This is essential to ensure that the dataset is free from errors, inconsistencies, or missing values, as these issues can significantly impact the accuracy and reliability of the subsequent analyses. Additionally, data organization and structuring are crucial for efficient data handling.  Furthermore, data visualization plays a pivotal role during these initial weeks. Visualizing data offers a unique perspective, allowing for the identification of trends, outliers, and patterns that might not be immediately evident through numerical analyses alone. Data visualization techniques, such as scatter plots, histograms, and heatmaps, will be employed to provide a graphical representation of the data, making it easier to comprehend and identify areas of interest.  The reason behind this comprehensive data exploration and visualization is to uncover intriguing insights and correlations within the dataset. These insights can serve as the foundation for the formulation of hypotheses and the design of machine learning algorithms. By understanding the data in-depth, it becomes possible to generate informed research questions that can be addressed through machine learning techniques.  The actual deployment of machine learning algorithms is scheduled to commence in December, following this crucial preparatory phase. By engaging in data exploration and visualization beforehand, the research team aims to enter the machine learning stage with a well-defined understanding of the data's characteristics and a clear sense of the questions to be answered. This sequential approach ensures that the machine learning algorithms are applied purposefully and effectively to derive meaningful results, enhancing the overall quality and impact of the research |

1. **Outline论文大纲**

|  |
| --- |
| I. Introduction  II. Literature Review  III. Data Collection and Integration  IV. Machine Learning in Asset Specificity and Performance of Outsourcing Contracts  V. Research Methodology |

1. **References 主要参考文献**

|  |
| --- |
| 1. How does offshore outsourcingof knowledge-intensive activities affect the exports and financial performance of emerging market firms? Peter J Buckley  2. The impact of knowledge flows on asset specificity from the perspective of open innovation - Feihu Zheng, Hao Jiao, Junyi Gu, Hwy-Chang Moon and Wenyan Yin  3. Reappraisal of outbound open innovation under the policy of China’s ‘Market for Technology’  Feihu Zheng, Hao Jiao & Hongbo Cai  4. Asset specificity and economic organization Author links open overlay panelMichael H. Riordan, Oliver E. Williamson  5. The Many Faces of Asset Specificity: A Critical Review of Key Theoretical Perspectives  Glauco De Vita, Arafet Tekaya, Catherine L. Wang  6. Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence PAUL L. JOSKOW  7. Asset specificity, firm heterogeneity and capital structure Srinivasan Balakrishnan, Isaac Fox  8. Asset specificity and firm value: Evidence from mergers Joon Ho Kim  9. Human asset specificity, physical asset specificity, and direct distribution Hidesuke Takata a, Mark E. Parry b  10. The Competitiveness of Networked Production: The Role of Trust and Asset Specificity Mick Carney  11. The role of asset specificity in the vertical integration decision Glen Whyte ∗  12. Asset Specificity of Nonfinancial Firms  Amir Kermani, Yueran Ma  13. Asset specificity roles in interfirm cooperation: Reducing opportunistic behavior or increasing cooperative behavior? S. Lui a, Yin-yee Wong b, Weiping Liu c  14. The contingent effects of asset specificity, contract specificity, and trust on offshore relationship performance Lei Wang a, Fuming Jiang b, Jun Li c, Kazuyuki Motohashi d, Xiaosong Zheng e  15. Asset specificity and the fear of exploitation Hwang  16. Asset specificity and relationship performance: A meta-analysis over three decades  Emanuela Delbufalo  17. Asset specificity and firm value: Evidence from mergers Joon Ho Kim  18. Asset specificity and conditional accounting conservatism Qingyuan Li, Li Xu  19. Does Governance Matter? Keiretsu Alliances and Asset Specificity as Sources of Japanese Competitive Advantage Jeffrey H. Dyer |

1. **Tutor’s Opinion 导师意见**

|  |
| --- |
| This drafted proposal put forths a meaningful research topic: The Impact of Asset Specificity on the Performance of Outsourcing Contracts, I believe it’s important while also challenging, so the author should read more literatures to clarity the key concepts’ definition, its influnencing mechanism and try more new analytical instrument including MLL, I believe the author can finish the thesis in the future.  **导师签名Tutor Signature：**  **2023 Year 11 Month 21 Day** |