

Phase_3_Project

README.md

1. Introduction: As a leading data mining and statistical computation company, Chris Tech specializes in transforming raw data into valuable insights. Signature Realtors, recognizing our expertise, has entrusted us with a strategic assignment. Our collaborative venture aims to develop a sophisticated price estimation model for houses, leveraging data extracted from jiji.com. This report delineates the meticulous approach we have adopted in the initial phases of the CRISP-DM framework to ensure the success of the project. Business Objectives: Our first objective in this collaboration is to create a precise price estimation model for houses, a task that aligns seamlessly with our core competencies in data mining and statistical computation. Through in-depth discussions with Signature Realtors, we've established clear success criteria, emphasizing accurate price predictions, model interpretability, and seamless integration into their existing systems. Our commitment to understanding the business objectives sets the stage for a purposeful and effective data mining project. The Situation: We recognize the importance of a comprehensive understanding of the project's context. Assessing the situation involves identifying the required resources, project requirements, potential risks, and conducting a cost-benefit analysis. We've meticulously evaluated the availability of human resources with expertise in data scraping and model development. Simultaneously, we've assessed the technological requirements and time commitments necessary for the successful completion of the project. Our risk assessment has identified potential challenges such as legal issues related to data scraping, data quality concerns, and model interpretability challenges. This detailed situational analysis, including a cost-benefit evaluation, ensures that our collaboration with Signature Realtors is grounded in a robust foundation. Data Mining Goals: The technical objectives of our data mining process have been clearly defined. These include specifying data cleaning processes, feature engineering techniques, model selection criteria, and the metrics by which we'll evaluate the success of our models. Additionally, we've outlined the data requirements, explicitly specifying the types of features needed for our price estimation model—factors such as size, bedrooms, bathrooms,

0

neighborhood characteristics, and country-specific attributes. Our meticulous approach to determining data mining goals ensures that we have a roadmap for technical success aligned with the overall business objectives. Project Plan: Selecting appropriate technologies and tools is paramount to the success of our project. We've chosen tools that facilitate efficient data scraping, robust data preprocessing, and streamlined model development. Our detailed project plan encompasses timelines and milestones for each phase of the project, providing clarity on tasks related to data scraping, preprocessing, model development, and evaluation. Additionally, we've established a communication plan to ensure seamless collaboration with Signature Realtors, fostering transparency and mutual understanding throughout the project lifecycle.

