1 Project Charta

1.1 Problem Definition

What is the problem?
The challenge is to develop a machine learning system that predicts the rental or selling price of residential properties. Currently, renters and buyers rely on manual and often subjective methods to assess property value. This leads to inefficiencies, missed opportunities, and mispriced listings.

Domain/Business Area
The project operates within the Swiss real estate market, involving rental agencies, private buyers and sellers, and expanding corporations. Real estate platforms like UrbanHome, ImmoScout24 and others are relevant sources.

Expected Benefit
The tool will help users, especially renters and buyers, find properties with the best price-to-quality ratio. It provides landlords and sellers with better pricing strategies and enables data-driven decision-making.

Why this effort?
An ML-powered property price predictor can dramatically reduce the time and uncertainty involved in the property search process. By learning from real data, the model helps stakeholders make informed choices aligned with their financial and lifestyle needs.

Stakeholders

Stakeholder	Role	Goal	Relationship
Buyers/Renters	Users of the predictor	Find the best value-for-money home	Interested in accuracy and usability
Sellers/Landlords	Input providers and users	Price fairly and competitively	Interested in fair predictions
Project Team	Developers and analysts	Build the system within 4 weeks	Provides value to both user groups
Data Sources	Real estate websites (UrbanHome etc.)	Provide structured/unstructured data	Data dependency

1.2 Situation Assessment

Resources
- **Personnel**: 3 people
- **Software/Tools**: VS Code, Python, SQLite, HTML, JS, CSS
- **Infrastructure**: Local development environments
- **Data Sources**: Real estate listing websites
- **Time**: 4 weeks (192 hours total)

Constraints
- All data must be acquired from public websites via scraping.
- Must handle dynamic content, inconsistent formats, and ensure data privacy and legality.

Restrictions
- Limited project time and small team size.
- No access to private or premium datasets.

Risks
- Lack of deep experience in ML model optimization.
- Time constraints may limit dashboard and DL integration quality.

1.3 Project Goals and Success Criteria

Qualitative Goals
- Develop a working ML model that predicts property prices.
- Visualize the results in a user-friendly dashboard.
- Enhance accuracy with tabular and image-based features.

Success Criteria and Metrics
- At least 1,000 unique properties scraped and processed.
- Classical regression model achieves MAE < 300 CHF or $R^2 > 0.75$.
- Dashboard allows users to query a property and receive a predicted price.
- Integration of deep learning features shows performance improvement over baseline.

Out of Scope
- Commercial deployment
- Legal contracts, compliance, and regulatory integration
- Mobile applications

1.4 Data Mining Goals

Task Mapping
- **Primary Task**: Regression, to predict continuous property prices
- **Subtasks**: - Classification, for image room types or amenity presence - Visualization, for dashboard output and EDA

Datasets to Be Used
- Scraped tabular data (price, rooms, size, location, etc.)
- Downloaded property images
- Optionally enriched with geographic/statistical data from APIs

Target Metrics
- **Regression**:
- MAE < 300 CHF
- $R^2 > 0.75$
- **Classification (images)**:
- Accuracy > 0.8
- F1-score > 0.75 (for multi-label classification)

1.5 Roles and Contact Details

Name	Role	Tasks	Contact
Joan	Data Engineer	Web scraping, data cleansing, DB design	dejondie@mail.gvsu.edu
Alessandro Mazzeo	ML Developer	Feature engineering, model training, evaluation	mazzale@students.zhaw.ch
Antonio Michal Verdile	Full Stack Dev	Dashboard development, ML integration, UI/UX design	verdiant@students.zhaw.ch