

ECOLE NATIONALE SUPÉRIEURE
D'INFORMATIQUE ET D'ANALYSE DES SYSTÈMES
UNIVERSITÉ MOHAMMED V

GÉNIE LOGICIEL
PROJET DE FIN DE DEUXIÈME ANNÉE

Simulateur de calcul des prix des
transactions immobilières

Réalisé par :
Tarmoune Oussama
Serroukh Anass



Encadré par :
W.Ettazi

2 juin 2024

Table des matières

1	Contexte et Objectfs	9
1.1	Contexte général	9
1.2	Problématique	9
1.3	Objectif	10
2	Méthodologie de développement et Planning du projet	11
2.1	Méthodologie de développement	11
2.2	Planning du projet	12
3	Le Web Scraping	13
3.1	Définition	13
3.2	Le Web Scraping dans notre Projet	13
3.3	Benchmark des Outils de Scraping	13
3.4	Méthodologie	14

Table des figures

2.1	Diagramme de Gantt	12
-----	------------------------------	----

Liste des tableaux

Remerciements

Nous tenons à exprimer notre profonde gratitude à notre professeur encadrante, Pr **Widad Ettazi**, pour son soutien constant et ses conseils avisés tout au long de ce projet. Votre expertise et votre disponibilité ont été des éléments clés dans la réussite de ce travail. Votre capacité à nous guider à travers les défis techniques et méthodologiques a été inestimable, et votre engagement envers notre réussite a été une source de motivation continue.

Résumé

Le projet **Simulateur de Calcul des Prix des Transactions Immobilières** vise à fournir une solution fiable et pratique pour estimer les prix de vente ou de location de biens immobiliers au Maroc. En raison du caractère privé des données immobilières dans le pays, nous avons opté pour le web scraping des sites d'annonces pour collecter les informations nécessaires. Après avoir nettoyé et structuré ces données, nous avons développé une application web permettant aux utilisateurs d'effectuer des estimations précises et de recevoir des notifications sur les évolutions des prix. Le site propose également un tableau de bord pour les administrateurs afin de surveiller et gérer les opérations du système.

Abstract

Real Estate Transaction Price Simulation project aims to provide a reliable and practical solution for estimating the sale or rental prices of real estate properties in Morocco. Given the private nature of real estate data in the country, we utilized web scraping techniques to gather necessary information from real estate listing websites. Following data collection, we implemented a thorough data cleaning process to ensure accuracy and reliability. The project culminated in the development of a web application that allows users to perform precise estimations and receive notifications about price changes. Additionally, the platform includes an administrative dashboard for monitoring and managing system operations. This project demonstrates the effective use of data scraping, cleaning, and modeling to deliver a valuable tool for the real estate market.

Mots clé

Immobilier, Estimation de Prix, Web Scraping, Nettoyage de Données, Modélisation de Données, MongoDB, Maroc, Annonces Immobilières, Système de Notifications, Tableau de Bord Administratif.

Keywords

Real Estate, Price Estimation, Web Scraping, Data Cleaning, Data Modeling, MongoDB, Morocco, Real Estate Listings, Notification System, Administrative Dashboard.

Introduction

Le marché immobilier au Maroc est en constante évolution, avec des prix de vente et de location fluctuants en fonction de divers facteurs économiques et géographiques. Cependant, l'accès à des données précises et actuelles sur les prix immobiliers reste un défi majeur, en raison de la nature privée de ces informations. Le projet **Simulateur de Calcul des Prix des Transactions Immobilières** a été conçu pour pallier cette lacune en fournissant un outil fiable permettant d'estimer les prix des biens immobiliers à partir de données réelles collectées en ligne.

Pour atteindre cet objectif, nous avons adopté une approche basée sur le web scraping, en ciblant les sites d'annonces immobilières tels qu'**avito.ma**. Cette méthode nous a permis de recueillir des données détaillées sur les propriétés, telles que les appartements, maisons, villas, bureaux et locaux commerciaux, selon les villes et le type de transaction (vente ou location). Ces données ont ensuite été nettoyées et structurées pour assurer leur précision et leur pertinence, permettant ainsi de développer une application web robuste offrant des estimations précises, des notifications sur les évolutions de prix, et un tableau de bord administratif pour la gestion des opérations.

Chapitre 1

Contexte et Objectfs

1.1 Contexte général

Le secteur immobilier joue un rôle crucial dans l'économie marocaine, représentant une part significative du PIB et influençant divers autres secteurs économiques. Toutefois, le marché immobilier marocain est caractérisé par un manque de transparence et d'accessibilité des données, ce qui complique la prise de décision pour les acheteurs, les vendeurs et les investisseurs. Les informations sur les prix des biens immobiliers sont souvent difficiles à obtenir, et les estimations disponibles peuvent varier considérablement, rendant la prévision des tendances du marché un défi de taille.

1.2 Problématique

La principale problématique dans le contexte du marché immobilier marocain réside dans l'accès limité à des données fiables et actuelles sur les prix des biens immobiliers. Cette absence de transparence crée une incertitude pour les acteurs du marché, qu'il s'agisse de particuliers souhaitant acheter ou vendre un bien, ou d'investisseurs cherchant à évaluer la rentabilité de leurs investissements. Les informations disponibles sont souvent fragmentées, obsolètes ou difficiles à comparer, ce qui rend la prise de décision complexe et risquée. De plus, les méthodes traditionnelles d'estimation des prix ne parviennent pas toujours à refléter les réalités du marché local, notamment en raison des variations régionales et des spécificités des biens immobiliers.

1.3 Objectif

Le projet **Simulateur de Calcul des Prix des Transactions Immobilières** a pour objectif de pallier cette problématique en développant une solution technologique capable de fournir des estimations de prix précises et fiables pour les biens immobiliers au Maroc. En utilisant des techniques avancées de web scraping, nous collectons des données détaillées et actualisées à partir de sites d'annonces immobilières. Ces données sont ensuite nettoyées, structurées et analysées pour offrir aux utilisateurs des estimations de prix basées sur des informations réelles et pertinentes. En plus de fournir des estimations, l'application propose des fonctionnalités telles que des notifications sur les évolutions de prix et un tableau de bord administratif pour la gestion et la surveillance des opérations.

L'objectif final est de créer un outil accessible et utile pour tous les acteurs du marché immobilier, facilitant ainsi la prise de décision et renforçant la confiance dans les transactions immobilières.

Chapitre 2

Méthodologie de développement et Planning du projet

2.1 Méthodologie de développement

Pour mener à bien le projet **Simulateur de Calcul des Prix des Transactions Immobilières**, nous avons adopté la méthodologie **Agile**. Cette approche itérative et flexible nous a permis de répondre de manière efficace aux exigences changeantes et d'améliorer continuellement le produit tout au long du cycle de développement. La méthodologie Agile se caractérise par des cycles de développement courts appelés "sprints", chacun d'une durée d'une semaine, au cours desquels des fonctionnalités spécifiques sont planifiées, développées, testées et évaluées.

1. **Analyse des Besoins et Définition des Données** : Comprendre les exigences du projet et définir les types de données nécessaires pour fournir des estimations précises.
2. **Collecte des Données via Web Scraping** : Récupérer les données nécessaires à partir des sites web d'annonces immobilières.
3. **Nettoyage et Structuration des Données** : Assurer que les données collectées sont précises, complètes et utilisables.
4. **Développement de l'Application Web** : Créer une interface utilisateur permettant d'accéder aux fonctionnalités d'estimation et de gestion des données.
5. **Modélisation et Analyse des Données** : Développer des modèles capables de fournir des estimations précises et des analyses statistiques.
6. **Déploiement et Automatisation** : Assurer que l'application est accessible et performante pour les utilisateurs finaux.

7. **Tests et Validation** : Garantir la fiabilité et la performance de l'application avant son lancement.

2.2 Planning du projet

Le diagramme de Gantt ci-dessous illustre la répartition des sprints et le calendrier prévu pour la réalisation du projet.

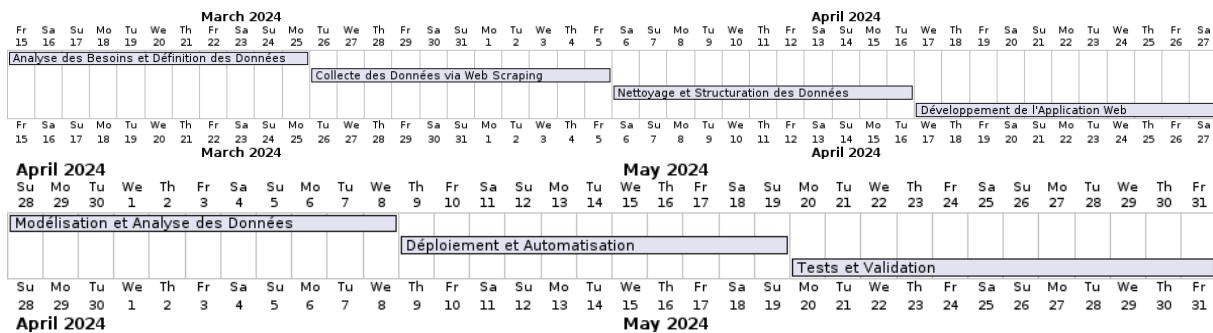


FIGURE 2.1 – Diagramme de Gantt

Chapitre 3

Le Web Scraping

3.1 Définition

Le web scraping, également appelé extraction de données web, est une technique utilisée pour extraire des informations à partir de sites web. Cette méthode implique l'utilisation de logiciels ou de scripts pour parcourir et récupérer des données disponibles publiquement sur internet. Contrairement aux API, qui fournissent des données de manière structurée et contrôlée, le web scraping permet d'extraire des informations directement à partir des pages web, en imitant la navigation et l'interaction d'un utilisateur humain.

3.2 Le Web Scraping dans notre Projet

Dans le cadre de notre projet Simulateur de Calcul des Prix des Transactions Immobilières, le web scraping a été essentiel pour collecter les données nécessaires en raison du caractère privé des informations immobilières au Maroc. Nous avons ciblé les sites d'annonces immobilières, notamment avito.ma, pour extraire les données relatives aux biens immobiliers (appartements, maisons, villas, bureaux, locaux commerciaux) en fonction des villes et du type de transaction (vente ou location). Cette approche nous a permis de constituer une base de données riche et détaillée pour notre application.

3.3 Benchmark des Outils de Scraping

Pour réaliser le web scraping, plusieurs outils et frameworks sont disponibles. Nous avons effectué un benchmark des outils les plus populaires pour

déterminer celui qui répondrait le mieux à nos besoins :

1. **BeautifulSoup** : Une bibliothèque Python utilisée pour parser des documents HTML et XML. Elle est simple d'utilisation mais nécessite d'être combinée avec d'autres bibliothèques pour la gestion des requêtes HTTP.
2. **Scrapy** : Un framework open-source en Python, spécialement conçu pour le web scraping. Il est très puissant et permet de gérer des projets de scraping complexes avec des fonctionnalités intégrées pour le crawling, le parsing, le stockage des données et la gestion des erreurs.
3. **Selenium** : Un outil permettant d'automatiser les navigateurs web. Il est utile pour scraper des sites dynamiques générés par JavaScript, mais il est plus lent et plus lourd que les autres options.
4. **Splash** : Un outil de rendu de pages web, capable d'exécuter JavaScript, qui peut être utilisé avec Scrapy pour scraper des sites dynamiques. Il offre une flexibilité pour gérer les pages complexes tout en restant intégré dans un environnement Scrapy.

Après avoir évalué ces options, nous avons opté pour Scrapy en combinaison avec Splash en raison de leur robustesse, de leur flexibilité et de leurs capacités avancées de gestion des projets de scraping.

3.4 Méthodologie

La méthodologie de web scraping que nous avons adoptée comprend les étapes suivantes :

1. Analyse des Sites Web Ciblés

- Identification des sites web pertinents pour le projet.
- Analyse de la structure HTML des pages pour déterminer les éléments contenant les informations nécessaires (titres, prix, descriptions, localisations, etc.).

2. Développement des Spiders

- Création de spiders avec Scrapy pour automatiser la navigation et l'extraction des données.
- Intégration de Splash pour gérer les pages dynamiques et exécuter JavaScript.

- Implémentation de mécanismes pour contourner les blocages de bots, tels que l'utilisation de proxies et de délais aléatoires.

3. Exécution et Collecte des Données

- Exécution des spiders pour collecter les données brutes.
- Surveillance des performances et des erreurs lors de l'extraction.

4. Nettoyage et Structuration des Données

- Nettoyage des données pour éliminer les doublons, les annonces fausses et compléter les informations manquantes.
- Conversion et structuration des données pour les rendre utilisables dans notre application.

5. Stockage des Données

- Stockage des données collectées et nettoyées dans une base de données.
- Mise en place de procédures de sauvegarde et de gestion des données.

En suivant cette méthodologie, nous avons pu obtenir une base de données fiable et complète, essentielle pour le fonctionnement de notre simulateur de calcul des prix des transactions immobilières.