

Lyra Immo – Fine-tuning d'un modèle GPT-3.5-turbo pour l'estimation immobilière

Objectif du projet

Créer un **assistant IA spécialisé dans l'estimation immobilière**, basé sur un modèle **GPT-3.5-turbo-0125 fine-tuné**, capable de fournir :

- Des **estimations cohérentes et directes**
 - Un **ton professionnel et stable**
 - Une **généralisation réaliste** sur des cas standards et liminaux
-

Étapes du projet

1. Préparation du dataset d'entraînement

- Création d'un fichier .jsonl de **150 exemples structurés**
- Paramètres utilisés :
 - Surface (m²)
 - Type de bien (T2, T3, villa...)
 - Nombre de pièces
 - État général
 - Localisation (quartier sensible, moyen, résidentiel, périurbain)
- Prix calculé selon le **barème Lyra Immo** :
 - Base : 2150 €/m²
 - Coefficients appliqués :
 - Quartier sensible : ×0.5
 - Quartier moyen ou résidentiel : ×1.4
 - Périurbain : ×1.6
 - Villa : ×1.25
 - Travaux : ×0.8

2. Création d'un dataset de validation

- Fichier .jsonl avec 30 exemples distincts
- Générés selon les mêmes règles
- Utilisé pour calculer la validation loss pendant l'entraînement

3. Fine-tuning OpenAI

- Modèle : gpt-3.5-turbo-0125
- Méthode : Supervised fine-tuning
- Suffixe : Lyra Immo
- Epochs : 3 (batch size 1, learning rate multiplier = 2)
- Validation loss agrégée finale : **0.917**
- Modèle sélectionné :
ft:gpt-3.5-turbo-0125:personal:lyra-immo:BXp5jzc7

Phase de test et validation

10 prompts ont été testés :

Cas standards (7 prompts) :

1. Appartement 82 m², bon état, quartier moyen → **246 000 €**
2. Maison 110 m², bon état, périurbaine → **429 000 €**
3. Villa 95 m², excellent état, résidentiel → **370 750 €**
4. Appartement 75 m², bon état, quartier moyen → **246 000 €**
5. Appartement 50 m², travaux, quartier sensible → **42 500 €**
6. Studio 45 m², quartier sensible → **38 250 €**
7. Appartement 60 m², T2, travaux, quartier sensible → **51 600 €**

Cas liminaux (3 prompts) :

8. Villa 180 m², excellent état, périurbaine → **702 000 €**
9. Studio 45 m², à rénover, quartier sensible → **42 500 €**
10. Appartement 50 m², travaux, quartier sensible → **42 500 €**

✅ Tous les cas ont été évalués comme **cohérents avec le barème Lyra Immo**, avec des écarts généralement **<10 %**, même en zone extrême.

Prompt 11 (supplémentaire) – : Test de robustesse linguistique (formulation utilisateur libre)

Prompt :

“Bonjour combien coûte une maison de 85 m² de plain-pied, en quartier pavillonnaire, avec 4 chambres et un jardin ?”

Ce test visait à évaluer la capacité du modèle fine-tuné à traiter une formulation **naturelle, non structurée**, typique d’un utilisateur non technicien.

Résultat du modèle fine-tuné :

“Bonjour, pour une maison de plain-pied de 85 m², avec 4 chambres et un jardin, dans un

quartier pavillonnaire, le prix estimé est 340 000 €. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”

✓ Le modèle :

- A reformulé proprement la demande,
- A généré une réponse dans le style narratif attendu,
- Et a fourni une **estimation chiffrée directement exploitable**, même si légèrement au-dessus du barème théorique.

Ce test démontre la **capacité de généralisation sémantique** de Lyra Immo, y compris **en dehors des prompts structurés** du jeu d'entraînement.

Exemples comparés – Réponses complètes des deux modèles

Cas	GPT-3.5 standard (non fine-tuné)	Lyra Immo fine-tuné
Appartement 82 m², bon état, quartier moyen	<i>“Bonjour, pour un appartement de 82 m², T3 en bon état, situé dans un quartier moyen, le prix estimé serait d'environ 180 000€ à 250 000€ en fonction de l'emplacement exact et des autres caractéristiques de l'appartement. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”</i>	<i>“Bonjour, pour un appartement de 82 m², en bon état, avec 3 pièces, en quartier moyen, le prix estimé est 246 000 €. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”</i>
Villa 180 m², excellent état, zone périurbaine	<i>“Bonjour, pour une villa de 180 m² en excellent état, située en zone périurbaine, le prix estimé serait d'environ 400 000€ à 500 000€ en fonction de l'emplacement exact et des autres caractéristiques de la propriété. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”</i>	<i>“Bonjour, pour une villa de 180 m², en excellent état, en zone périurbaine, le prix estimé est 702 000 €. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”</i>
Studio 45 m², à rénover, quartier sensible	<i>“Bonjour, pour un studio de 45 m² nécessitant des travaux, situé dans un quartier sensible, le prix estimé serait d'environ 30 000€ à 60 000€ en fonction de l'étendue des travaux à réaliser et des caractéristiques du studio. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”</i>	<i>“Bonjour, pour un studio de 45 m², à rénover, en quartier sensible, le prix estimé est 38 250 €. Nous restons à votre disposition pour tout besoin d'informations complémentaire. Avec toute notre considération. Lyra Immo.”</i>

🌟 Points forts du modèle fine-tuné

- Réponse **chiffrée directe**, sans fourchette
- **Ton professionnel**, fluide et constant
- Structure de réponse complète avec introduction et formule de courtoisie
- **Capacité à traiter les cas rares ou extrêmes (biens liminaux)** avec stabilité
- Temps de réponse court, intégrable dans Make ou tout outil d'agentisation

📊 Perspectives d'évolution

+ Enrichissement par données synthétiques encadrées

- Pour corriger un **léger déséquilibre** observé sur certains biens haut de gamme
- Ajout de cas rares (ex. : villas > 700 000 €, studios rénovés en centre-ville)
- Création possible d'un **dataset complémentaire V2**, intégrant :
 - Données fictives cohérentes validées par un expert ou une logique barémique
 - Reformulations naturelles pour robustesse linguistique