|  |
| --- |
| **BỘ GIÁO DỤC VÀ ĐÀO TẠO**  **TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HỒ CHÍ MINH** |
| **KHOA CÔNG NGHỆ THÔNG TIN**  **'Y&&Y'**  **Logo ĐH Sư Phạm Kỹ Thuật TP Hồ Chí Minh-HCMUTE** |

|  |
| --- |
| **BÁO CÁO ĐỒ ÁN CUỐI KỲ**  **ĐỀ TÀI: HOUSE RENT PRICING** |
| **Môn học: KHAI PHÁ DỮ LIỆU**  **Mã lớp học phần: 222DAMI330484**  **GVHD: ThS. Nguyễn Văn Thành** |
| **Nhóm sinh viên thực hiện: Nhóm 11**   |  |  | | --- | --- | | Trần Lê Ngọc Gia Hân | 20133038 | | Nguyễn Thị Hoàng Trang | 20133099 | | Nguyễn Thị Thanh Ngân | 20133018 | |

**TP. Hồ Chí Minh, 13 tháng 05 năm 2023**

**DANH SÁCH THÀNH VIÊN THAM GIA**

**THỰC HIỆN ĐỀ TÀI VÀ VIẾT BÁO CÁO**

**Môn: Khai Phá Dữ Liệu -** *HỌC KÌ II – NĂM HỌC 2022 – 2023*

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **HỌ VÀ TÊN** | **MSSV** | **TỶ LỆ ĐÓNG GÓP** |
| 1 | Trần Lê Ngọc Gia Hân | 20133038 | 100% |
| 2 | Nguyễn Thị Hoàng Trang | 20133099 | 100% |
| 3 | Nguyễn Thị Thanh Ngân | 20133018 | 100% |

**Nhận xét của giảng viên:**

**……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………**

*Ngày … tháng 05 năm 2023*

*Giảng viên chấm điểm*

*Ths. Nguyễn Văn Thành*

**LỜI CẢM ƠN**

Lời đầu tiên, nhóm chúng em xin gửi lời cảm ơn chân thành đến **Trường Đại học Sư phạm Kỹ thuật TPHCM** đã đưa môn học **Khai Phá Dữ Liệu** vào chương trình giảng dạy. Đặc biệt, chúng em xin gửi lời cảm ơn sâu sắc đến giảng viên bộ môn – **Th.S Thầy Nguyễn Văn Thành** đã dạy dỗ, truyền đạt những kiến thức quý báu cho chúng em trong suốt thời gian học tập vừa qua. Trong thời gian tham gia lớp học Kho Dữ Liệu của Thầy, nhóm em đã có thêm cho mình nhiều kiến thức bổ ích, tinh thần học tập hiệu quả, nghiêm túc. Đây chắc chắn sẽ là những kiến thức quý báu, là hành trang để chúng em có thể vững bước sau này. Bộ môn Khai Phá Dữ Liệu là môn học thú vị, vô cùng bổ ích và có tính thực tế cao. Đảm bảo cung cấp đủ kiến thức, gắn liền với nhu cầu thực tiễn của sinh viên. Mặc dù chúng em đã cố gắng hết sức nhưng chắc chắn bài đồ án này khó có thể tránh khỏi những thiếu sót và nhiều chỗ còn chưa chính xác, kính mong Thầy xem xét và góp ý để bài đồ án của chúng em được hoàn thiện hơn.

Nhóm chúng em xin chân thành cảm ơn!

**MỤC LỤC**

**PHÂN CÔNG NHIỆM VỤ**

# CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI

## Lý do chọn đề tài

* Nhà ở ở Ấn Độ rất đa dạng, từ cung điện của các maharaja trước đây đến các tòa nhà chung cư hiện đại ở các thành phố lớn đến những túp lều nhỏ ở những ngôi làng xa xôi. Đã có sự tăng trưởng to lớn trong lĩnh vực nhà ở của Ấn Độ khi thu nhập tăng lên.

## Tổng quan về tập dữ liệu

### Nguồn dữ liệu

* Nhóm sử dụng tập dữ liệu **House Rent Dataset** trên Kaggle. Dữ liệu này cung cấp thông tin về các căn hộ/phòng cho thuê ở Ấn Độ được mô tả dựa trên nhiều đặc điểm khác nhau như kích thước phòng, vị trí căn hộ, số lượng phòng cho thuê, đối tượng ưu tiên,.. giá cho thuê.
* Nguồn dữ liệu: [House Rent Dataset Kaggle](https://www.kaggle.com/datasets/iamsouravbanerjee/house-rent-prediction-dataset?select=House_Rent_Dataset.csv&fbclid=IwAR3UvrpfqqCkZzZs9bwekZ_dAggQrQVAxoLnKjFVeC0a-c-5B1MABpmpn_Y)

### Mô tả chi tiết tập dữ liệu

* Tập dữ liệu gồm có hơn 4.700 dòng (rows) là các nhà cho thuê và 12 cột (columns) là thông tin về căn hộ cho thuê đó. Cụ thể là
* BHK: Số lượng phòng Bedrooms, Hall, Kitchen
* Rent: Giá cho thuê
* Size: Kích thước (m)
* Floor: Houses/Apartments/Flats situated in which Floor and Total Number of Floors (Example: Ground out of 2, 3 out of 5, etc.)
* Area Type: Size of the Houses/Apartments/Flats calculated on either Super Area or Carpet Area or Build Area.
* Area Locality: Locality of the Houses/Apartments/Flats.
* City: City where the Houses/Apartments/Flats are Located.
* Furnishing Status: Furnishing Status of the Houses/Apartments/Flats, either it is Furnished or Semi-Furnished or Unfurnished.
* Tenant Preferred: Type of Tenant Preferred by the Owner or Agent.
* Bathroom: Number of Bathrooms.
* Point of Contact: Whom should you contact for more information regarding the Houses/Apartments/Flats.