**TRƯỜNG ĐẠI HỌC THỦY LỢI**

**A blue and white logo

AI-generated content may be incorrect.**

**BÁO CÁO THỰC HÀNH 1**

**HỌC PHẦN TIME SERIES**

**Hà Nội, năm 2025**

**A.Tổng quan về dữ liệu và chuẩn bị dữ liệu**

**I.Tổng quan**

**A screenshot of a computer

AI-generated content may be incorrect.**

**File với 68MB gồm 12 cột và 750000 dòng dữ liệu:  
id, Podcast\_Name ; Episode\_Title; Episode\_Length\_minutes; Genre; Host\_Popularity\_percentage; Publication\_Day; Publication\_Time; Guest\_Popularity\_percentage; Number\_of\_Ads; Episode\_Sentiment; Listening\_Time\_minutes;**

**II.Chuẩn bị dữ liệu thực hành**

**1.Xử lý dữ liệu**

**Đề bài: 2251262644 %7 +1= 2 -> đề 2: Tuesday**

**Sau khi tải file dữ liệu trên Kaggle em tiến hành lọc dữ liệu nhằm lấy những hàng với Tuesday để chạy mô hình.**

**A screenshot of a computer screen

AI-generated content may be incorrect.**

H1.Code lọc lấy những hàng Tuesday trong Publication\_Day

-Kiểm tra dữ liệu:

A screenshot of a computer

AI-generated content may be incorrect.

H2. Số lượng dữ liệu thiếu

-Thấy được dữ liệu thiếu ở 2 cột :  
+Episode\_Length\_minutes

+Guest\_Popularity\_percentage

-> Thực hiện xử thiếu với giá trị trung bình



\*Kiểm tra lại dữ liệu:

A screenshot of a computer

AI-generated content may be incorrect.

**B.Khởi chạy mô hình với Kalman**

**I. GIỚI THIỆU**

Trong bài toán phân tích dữ liệu chuỗi thời gian như lượt xem chương trình truyền hình , việc làm mịn (smoothing) dữ liệu và phát hiện xu hướng là rất quan trọng để hỗ trợ ra quyết định trong sản xuất và quảng bá nội dung.

Báo cáo này trình bày việc áp dụng 3 mô hình Kalman Filter vào bộ dữ liệu về lượt xem của các chương trình truyền hình:

1. Kalman Filter (KF) – Mô hình xác suất tuyến tính.
2. Normalizing Kalman Filter (NK) – KF kết hợp chuẩn hóa dữ liệu đầu vào.
3. Deep Kalman Filter (DKF) – Mô hình học sâu để học trạng thái ẩn phi tuyến.

**II. CƠ SỞ LÝ THUYẾT**

**1. Kalman Filter (KF)**

Là một thuật toán đệ quy dùng để ước lượng trạng thái của hệ thống động có nhiễu Gaussian, phù hợp với chuỗi thời gian ổn định và tuyến tính.

**2. Normalizing Kalman Filter (NK)**

Mở rộng của KF bằng cách chuẩn hóa dữ liệu trước khi áp dụng mô hình nhằm tăng độ ổn định và khả năng tổng quát.

**3. Deep Kalman Filter (DKF)**

Kết hợp mạng nơ-ron với mô hình Kalman để học ánh xạ phi tuyến giữa quan sát và trạng thái ẩn, phù hợp với dữ liệu phức tạp.

**III. PHƯƠNG PHÁP THỰC HIỆN**

**1. Chọn dữ liệu**

* Gom nhóm dữ liệu theo tuần (**Week**) → lấy trung bình lượt xem mỗi tuần.

**2. Triển khai mô hình**

**a. Mô hình 1: Kalman Filter**

**b. Mô hình 2: Normalizing Kalman Filter**

**c. Mô hình 3: Deep Kalman Filter**

\*Biểu đồ so sánh 3 mô hình:  
A graph with lines and numbers

AI-generated content may be incorrect.

* KF : Hiệu quả tốt với dữ liệu ít biến động, nhưng phản ứng chậm với thay đổi đột ngột.
* NK : Cải thiện đáng kể khả năng lọc nhiễu nhờ chuẩn hóa.
* DKF : Học được xu hướng dài hạn, nhưng cần nhiều dữ liệu và thời gian huấn luyện.