**ĐẠI HỌC QUỐC GIA HÀ NỘI**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

**Nguyễn Ngọc Thái**

**NGHIÊN CỨU BÀI TOÁN PHÁT HIỆN ĐỘNG ĐẤT SỬ DỤNG DỮ LIỆU CẢM BIẾN GIA TỐC**

Ngành: Công nghệ Kỹ thuật Cơ điện tử

**TÓM TẮT KHOÁ LUẬN TỐT NGHIỆP**

**HÀ NỘI - 2025**

**TÓM TẮT**

Với sự tiến bộ vượt bậc của khoa học và công nghệ, đặc biệt là trong lĩnh vực trí tuệ nhân tạo (AI) và Internet vạn vật (IoT), việc triển khai các hệ thống tự động hóa phát hiện sớm thiên tai đang trở thành xu thế tất yếu trong công tác phòng chống và giảm nhẹ rủi ro. Trong đó, phát hiện sớm động đất là một trong những ứng dụng quan trọng và cấp thiết nhất. Bên cạnh công nghệ cảm biến hiện đại cho phép thu thập dữ liệu liên tục và chính xác từ hàng nghìn điểm quan trắc. Khi được kết hợp với AI và học máy, các hệ thống này không chỉ phát hiện rung chấn bất thường trong thời gian thực, mà còn có thể phân biệt được giữa động đất thật và nhiễu nền, việc đưa ra cảnh báo sớm chỉ trong vài giây điều có thể quyết định giữa sự sống và cái chết cho hàng ngàn người.

Đồ án tập trung nghiên cứu và xây dựng mô hình phát hiện động đất dựa trên dữ liệu từ cảm biến gia tốc ba trục. Trong bối cảnh Việt Nam và nhiều quốc gia khác ngày càng đối mặt với rủi ro thiên tai từ động đất, việc phát hiện sớm các tín hiệu địa chấn là một yêu cầu cấp thiết nhằm giảm thiểu thiệt hại về người và tài sản. Dữ liệu được sử dụng chủ yếu từ hai nguồn: K-NET (Nhật Bản) và Italia, bao gồm các tín hiệu động đất thật và nhiễu nền như xe cộ, gió,... Sau khi thu thập và trải qua các quá trình tiền xử lý và quá trình trích xuất đặc trưng từ tín hiệu. Đồ án triển khai trên nhiều mô hình học máy ML như Random Forest, SVM, Logistic Regression,... với quy trình huấn luyện, đánh giá theo chuẩn scikit-learn để chọn ra mô hình đánh giá tốt nhất.

Kết quả cho thấy mô hình có thể nhận diện hiệu quả tín hiệu động đất so với nhiễu với độ chính xác cao. Tuy nhiên, đề tài hiện vẫn chỉ dừng ở mức mô phỏng offline và chỉ có chức năng phân biệt giữa tín hiệu động đất và nhiễu chứ chưa áp dụng real-time và chưa tích hợp vào mạng cảnh báo sớm. Từ đó, đồ án đề xuất định hướng tiếp theo là phát triển ứng dụng thời gian thực, mở rộng dữ liệu địa phương và tăng cường học sâu để cải thiện hiệu năng.

Dưới sự hướng dẫn của thầy Hoàng Văn Mạnh và sự đồng ý của Khoa Cơ kỹ thuật và Tự động hóa, em đã thực hiện chuyên đề “Nghiên cứu bài toán phát hiện động đất sử dụng dữ liệu cảm biến gia tốc”.

## Tính cấp thiết đề tài

Động đất, một trong những hiện tượng tự nhiên có sức tàn phá khủng khiếp nhất, luôn là mối đe dọa thường trực đối với sự an toàn của con người và sự ổn định của xã hội. Sự rung chuyển đột ngột của vỏ Trái Đất không chỉ gây ra những tổn thất nặng nề về người và tài sản mà còn để lại những hậu quả lâu dài về mặt kinh tế, xã hội và môi trường. Động đất không chỉ gây ra những thiệt hại trực tiếp về người và tài sản mà còn ảnh hưởng sâu sắc đến các hoạt động kinh tế, xã hội. Các công trình giao thông, hệ thống điện, nước, các cơ sở y tế và giáo dục có thể bị phá hủy, gây ra sự gián đoạn trong hoạt động sản xuất và sinh hoạt của người dân. Hơn nữa, động đất còn gây ra những tác động tiêu cực đến môi trường như sạt lở đất, lũ quét, ô nhiễm nguồn nước, ảnh hưởng đến hệ sinh thái và đa dạng sinh học. Theo Ngân hàng Thế giới, thiệt hại do thiên tai, trong đó có động đất, có thể làm giảm GDP của một quốc gia tới 1% mỗi năm.

Hình trên đây thể hiện sự phân bố các trận động đất có độ lớn trên 4.0 xảy ra trên toàn thế giới trong năm 2015. Có thể thấy, phần lớn các trận động đất tập trung dọc theo các ranh giới mảng kiến tạo, đặc biệt là khu vực bao quanh Thái Bình Dương và trải dài qua Nhật Bản, Indonesia, Philippines, New Zealand, và bờ Tây châu Mỹ. Ngoài ra, các vùng như Trung Á, Iran, Thổ Nhĩ Kỳ và rìa Địa Trung Hải cũng ghi nhận nhiều hoạt động địa chấn. Ngược lại, châu Phi, Bắc Âu và phần lớn nội địa châu Á ít xuất hiện các trận động đất đáng kể.

**Tình hình Động đất trên thế giới**

Theo thống kê từ Cục Khảo sát Địa chất Hoa Kỳ (USGS), mỗi năm trên thế giới xảy ra khoảng 20.000 trận động đất, trong đó có khoảng 18 trận động đất mạnh từ 7.0 độ Richter trở lên. Chỉ tính riêng trong thế kỷ 21, đã có hàng triệu người thiệt mạng do động đất. Trận động đất Tohoku năm 2011 tại Nhật Bản, với cường độ 9.0 độ Richter, đã gây ra sóng thần tàn phá nặng nề, cướp đi sinh mạng của hơn 18.000 người và gây thiệt hại kinh tế ước tính lên đến 360 tỷ USD. Hay trận động đất năm 2010 tại Haiti, dù có cường độ thấp hơn (7.0 độ Richter), nhưng do điều kiện xây dựng kém, đã làm hơn 200.000 người thiệt mạng. Những con số này là minh chứng rõ ràng cho sức tàn phá khủng khiếp của động đất và tầm quan trọng của việc cảnh báo sớm.

**Tình hình Động đất tại Việt Nam**

Việt Nam tuy không nằm trên vành đai lửa Thái Bình Dương – khu vực có hoạt động địa chấn mạnh nhất thế giới, nhưng vẫn chịu ảnh hưởng bởi các hoạt động kiến tạo địa chất nội địa và khu vực lân cận. Các trận động đất tại Việt Nam chủ yếu liên quan đến hệ thống đứt gãy kiến tạo đang hoạt động. Đặc biệt, khu vực Tây Bắc với các đứt gãy như Điện Biên – Mường Lay, Sông Mã – Tuần Giáo – Lai Châu đã từng ghi nhận những trận động đất mạnh.

Theo báo điện tử Tiền Phong, đưa tin vào ngày 29/03/2025 về “Những trận động đất lớn nhất ở Việt Nam thế kỷ qua” có thể kể đến như sau:

* Ngày 1/11/1935, một trận động đất 6.8 độ richter xảy ra tại Điện Biên.
* Năm 1983 tại Tuần Giáo (Điện Biên) cũng trải qua trận động đất 6,7 độ richter, gây thiệt hại đáng kể về cơ sở hạ tầng và kinh tế.
* Gần đây nhất, vào ngày 28/7/2024, liên tục nhiều trận động đất 5 độ đã gây ra rung chấn cho khắp Tây Nguyên và nhiều tỉnh miền Trung.

Ngoài ra, những hoạt động nhân sinh cũng góp phần làm kích thích động đất. Việc xây dựng và vận hành các hồ chứa thủy điện lớn làm thay đổi áp lực nước và ứng suất trong lòng đất, có thể kích hoạt các đứt gãy tiềm ẩn. Tại huyện Kon Plông, tỉnh Kon Tum, sau khi các hồ chứa được tích nước, khu vực này đã ghi nhận nhiều trận động đất liên tiếp, cho thấy mối liên hệ giữa hoạt động nhân sinh và động đất kích thích. ​

Từ yếu tố trên cho chúng ta thấy, dù không nằm trên vành đai lửa, Việt Nam vẫn đối mặt với nguy cơ động đất từ cả hoạt động kiến tạo tự nhiên và tác động nhân sinh. Việc nghiên cứu, theo dõi và đánh giá các đứt gãy hoạt động, cùng với quản lý chặt chẽ các hoạt động có thể kích thích động đất, là cần thiết để giảm thiểu rủi ro và thiệt hại do động đất gây ra.

## Ý nghĩa khoa học và thực tiễn

Trong những năm gần đây, sự phát triển của công nghệ cảm biến, đặc biệt là cảm biến gia tốc, đã mở ra những hướng đi mới trong việc phát hiện động đất. Cảm biến gia tốc có khả năng ghi nhận những dao động nhỏ nhất của mặt đất, từ đó cung cấp dữ liệu quan trọng để phân tích và nhận diện dấu hiệu của động đất. Với sự phổ biến của các thiết bị di động thông minh, việc tích hợp cảm biến gia tốc vào hệ thống cảnh báo động đất có thể tạo ra một mạng lưới giám sát rộng khắp, giúp chúng ta phát hiện động đất một cách nhanh chóng và chính xác hơn.

Đề tài "Phát hiện động đất sử dụng dữ liệu cảm biến gia tốc" của em tập trung vào việc nghiên cứu khả năng ứng dụng của cảm biến gia tốc trong việc phát hiện động đất. Việc nghiên cứu đề tài này không chỉ có ý nghĩa khoa học mà còn mang tính ứng dụng thực tiễn cao, góp phần vào việc nâng cao khả năng cảnh báo và giảm thiểu thiệt hại do động đất gây ra.

Hệ thống này hướng đến việc hỗ trợ phát hiện sớm các trận động đất một cách tự động, nhằm nâng cao hiệu quả cảnh báo và giảm thiểu rủi ro thiên tai.Thông qua đó, nghiên cứu kỳ vọng đóng góp một phần nhỏ vào việc tự động hóa quá trình nhận diện tín hiệu địa chấn, giúp tăng độ chính xác và rút ngắn thời gian cảnh báo so với các phương pháp truyền thống

## Đối tượng và phương pháp nghiên cứu

Đề tài chỉ phân loại tín hiệu thành hai nhóm: Động đất (EQ) và Nhiễu (Noise) và có phân tích thêm độ lớn độ lớn chứ không phân tích thêm các thông tin như tọa độ tâm chấn, thời gian xảy ra hoặc mức độ ảnh hưởng.

Dữ liệu huấn luyện và kiểm thử mang tính giới hạn về số lượng và khu vực địa lý, chủ yếu thu thập được từ các nguồn mô phỏng hoặc tập dữ liệu công khai. Dữ liệu trong quá trình thực hiện chủ yếu sử dụng hai bộ dataset đó là bộ **K-NET** của Cơ quan Khí tượng Nhật Bản và bộ thứ hai là bộ dataset về **động đất của Italia.**

## Nội dung nghiên cứu

Quá trình nghiên cứu tập trung vào các mục tiêu chính như sau:

1. Thu thập và tiền xử lý dữ liệu: Chuyển đổi đơn vị đo, loại bỏ nhiễu và giá trị ngoại lai, chuẩn hóa dữ liệu và cân bằng dữ liệu để đảm bảo đầu vào phù hợp với các mô hình học máy.
2. Trích xuất đặc trưng: Tính toán các đặc trưng thống kê và phổ tần số như: năng lượng, IQR, ZC, PSD, kurtosis, skew,... từ tín hiệu gia tốc theo từng trục.
3. Xây dựng và huấn luyện mô hình: Sử dụng các thuật toán học máy như Random Forest, SVM, và Logistic Regression. Mô hình được huấn luyện với tập dữ liệu được chia theo kỹ thuật cross-validation để đảm bảo độ chính xác và tính tổng quát.
4. Đánh giá hiệu quả mô hình: Dựa trên các tiêu chí như Accuracy, Precision, Recall, F1-score và AUC-ROC để so sánh hiệu suất phân loại động đất và nhiễu.
5. Triển khai giao diện đơn giản (dạng phần mềm hoặc ứng dụng mô phỏng) cho phép người dùng nạp dữ liệu và nhận kết quả phân loại: “Động đất” hay “Nhiễu”

Một trong những giới hạn quan trọng của đề tài là chưa triển khai hệ thống phát hiện động đất theo thời gian thực (real-time), cũng như chưa tích hợp vào một mạng lưới cảnh báo sớm hoàn chỉnh. Các mô hình được xây dựng và đánh giá chủ yếu trong môi trường thử nghiệm, với dữ liệu được xử lý sẵn và thực thi dưới dạng mô phỏng. Điều này đồng nghĩa với việc kết quả nghiên cứu hiện tại chỉ mới kiểm chứng được khả năng phân loại tín hiệu trong điều kiện lý tưởng, chưa phản ánh đầy đủ hiệu quả khi áp dụng trong điều kiện thực tế.