

## LỜI CAM ĐOAN

Trên thực tế không có sự thành công nào mà không gắn liền với những sự hỗ trợ, giúp đỡ dù ít hay nhiều, dù trực tiếp hay gián tiếp của người khác. Trong suốt thời gian từ khi bắt đầu học tập tại trường đến nay, tôi đã nhận được rất nhiều sự quan tâm, giúp đỡ của quý Thầy (Cô), gia đình và bạn bè. Với lòng biết ơn sâu sắc nhất, tôi xin gửi đến quý Thầy (Cô) **Ở Khoa Công Nghệ Thông Tin Trường Đại Học Điện Lực** đã cùng với tri thức và tâm huyết của mình để truyền đạt vốn kiến thức quý báu cho tôi trong suốt thời gian học tập tại trường. Và đặc biệt, trong thời gian thực hiện luận văn này tôi cũng xin chân thành cảm ơn **TS. Nguyễn Thị Thu Hà** đã nhiệt tình hướng dẫn, nếu không có những lời hướng dẫn của Cô thì bài luận văn này rất khó có thể hoàn thiện được. Một lần nữa, tôi xin chân thành cảm ơn Cô.

Tuy nhiên do thời gian và nguồn lực có hạn nên không tránh khỏi những thiếu sót. Rất mong nhận được những ý kiến đóng góp quý báu từ quý Thầy Cô để kiến thức của tôi trong lĩnh vực này được hoàn thiện hơn.

Người thực hiện đề tài

**Vũ Đức Khánh**

## MỤC LỤC

|                                                                                                |           |
|------------------------------------------------------------------------------------------------|-----------|
| <b>CHƯƠNG 1: TỔNG QUAN VỀ ĐIỆN MẶT TRỜI.....</b>                                               | <b>1</b>  |
| 1.1 Giới thiệu .....                                                                           | 1         |
| 1.1.1 Khái niệm về điện mặt trời.....                                                          | 1         |
| 1.1.2 Nguyên tắc cơ bản của bức xạ mặt trời .....                                              | 1         |
| 1.2 Một số hệ thống điện mặt trời phổ biến .....                                               | 5         |
| 1.3 Một số phương pháp dự báo .....                                                            | 6         |
| 1.3.1 Phương pháp thống kê kinh tế .....                                                       | 6         |
| 1.4 Phương pháp đánh giá dự báo.....                                                           | 10        |
| 1.5 Kết luận chương 1 .....                                                                    | 11        |
| <b>CHƯƠNG 2: DỰ BÁO NHU CẦU SỬ DỤNG ĐIỆN MẶT TRỜI.....</b>                                     | <b>12</b> |
| <b>SỬ DỤNG PHƯƠNG PHÁP HỌC MÁY .....</b>                                                       | <b>12</b> |
| 2.1 Một số phương pháp liên quan .....                                                         | 12        |
| 2.1.1 Multi-agent models .....                                                                 | 13        |
| 2.1.2 Các mô hình cơ bản .....                                                                 | 14        |
| 2.1.3 Mô hình reduced-form .....                                                               | 15        |
| 2.1.4 Mô hình thống kê và kinh tế lượng .....                                                  | 15        |
| 2.1.5 Tính toán thông minh.....                                                                | 17        |
| 2.1.5.1 Mạng Nơ-ron nhân tạo .....                                                             | 17        |
| 2.1.5.2 Máy Vec-tơ hỗ trợ (SVMs) .....                                                         | 18        |
| 2.1.5.3 Suy diễn mờ.....                                                                       | 18        |
| 2.1.5.4 Mạng nơ-ron học sâu.....                                                               | 19        |
| 2.2 Dự báo nhu cầu sử dụng điện mặt trời dựa trên học máy .....                                | 20        |
| 2.2.1 Mô hình bài toán.....                                                                    | 20        |
| 2.2.2 Phân tích các yếu tố ảnh hưởng .....                                                     | 21        |
| 2.3 Ứng dụng máy Vec-tơ hỗ trợ (SVMs) trong bài toán dự báo nhu cầu sử dụng điện mặt trời..... | 28        |
| 2.3.1 Máy Vec-tơ hỗ trợ (SVMs) .....                                                           | 28        |
| 2.3.2 Phương pháp VMS tuyến tính .....                                                         | 31        |
| 2.3.3 Phân tách tuyến tính theo không gian đầu vào.....                                        | 32        |

|                                                                               |           |
|-------------------------------------------------------------------------------|-----------|
| 2.3.4 Phương pháp VMS phi tuyến tính.....                                     | 36        |
| 2.3.5 Đánh giá và phân tích kết quả.....                                      | 37        |
| 2.4 Kết luận chương 2.....                                                    | 38        |
| <b>CHƯƠNG 3: XÂY DỰNG MÔ HÌNH DỰ BÁO TÌNH HÌNH SỬ DỤNG ĐIỆN MẶT TRỜI.....</b> | <b>39</b> |
| 3.1 Môi trường cài đặt.....                                                   | 39        |
| 3.2 Ngôn ngữ sử dụng.....                                                     | 39        |
| 3.3 Kiến trúc hệ thống.....                                                   | 40        |
| 3.4 Quy trình thực nghiệm.....                                                | 43        |
| 3.5 Kết quả thử nghiệm.....                                                   | 44        |
| 3.5.1 Nguồn dữ liệu sử dụng.....                                              | 44        |
| 3.5.2 Dữ liệu dùng cho thử nghiệm.....                                        | 45        |
| 3.6 Một số giao diện của chương trình.....                                    | 46        |
| 3.7 Kết luận chương 3.....                                                    | 47        |
| <b>KẾT LUẬN.....</b>                                                          | <b>48</b> |
| <b>TÀI LIỆU THAM KHẢO.....</b>                                                | <b>50</b> |

## DANH MỤC BẢNG BIỂU

|                                                                  |    |
|------------------------------------------------------------------|----|
| Bảng 1.1: Tỷ lệ phần trăm thu nhập và tiêu dùng thời trang ..... | 7  |
| Bảng 2.1: Kết quả cho các mô hình học tập khác nhau .....        | 19 |
| Bảng 3.1: Dữ liệu huấn luyện .....                               | 42 |
| Bảng 3.2: Lưu dự báo qua các lần thử nghiệm .....                | 43 |
| Bảng 3.3: Dữ liệu thống kê về khách hàng và sản lượng phát ..... | 45 |

## DANH MỤC HÌNH VẼ

|                                                                                                                                             |    |
|---------------------------------------------------------------------------------------------------------------------------------------------|----|
| Hình 1.1: Mô hình lương và chi tiêu trong gia đình. ....                                                                                    | 6  |
| Hình 1.2: Kết quả thử nghiệm với mô hình HW .....                                                                                           | 9  |
| Hình 1.3: Kết quả thử nghiệm với mô hình ARIMA.....                                                                                         | 9  |
| Hình 2.1: Một số phương pháp tiếp cận dự báo.....                                                                                           | 12 |
| Hình 2.2: Phân nhóm các phương pháp dự báo nhu cầu sử dụng điện.....                                                                        | 13 |
| Hình 2.4: Mô hình hệ thống điều khiển mờ.....                                                                                               | 18 |
| Hình 2.5: Dự báo tải với mạng thần kinh tái diễn.....                                                                                       | 19 |
| Hình 2.8: Ví dụ minh họa về phân loại k- láng giềng gần nhất. ....                                                                          | 30 |
| Hình 2.9: Nguồn gốc của một máy vector hỗ trợ tuyến tính cho (a) có thể chia tách tuyến tính và (b) các điểm dữ liệu không tuyến tính. .... | 33 |
| Hình 3.1: Thành phần chính của Microsoft .NET .....                                                                                         | 39 |
| Hình 3.2: Microsoft .NET Framework .....                                                                                                    | 40 |
| Hình 3.3: Sơ đồ chức năng hệ thống dự báo.....                                                                                              | 41 |
| Hình 3.4: Biểu đồ Use case tổng quát.....                                                                                                   | 42 |
| Hình 3.5: Tập đoàn Điện lực Việt Nam.....                                                                                                   | 45 |
| Hình 3.6: Tổng cục thống kê Việt Nam.....                                                                                                   | 45 |
| Hình 3.7: Giao diện dữ liệu huấn luyện cho chương trình .....                                                                               | 46 |
| Hình 3.8: Giao diện dự báo nhu cầu sử dụng điện mặt trời.....                                                                               | 47 |

## DANH MỤC TỪ VIẾT TẮT

| STT | Từ viết tắt | Nội dung                    |
|-----|-------------|-----------------------------|
| 1   | NLMT        | Năng lượng mặt trời         |
| 2   | ANN         | Mạng nơ ron nhân tạo        |
| 3   | DNN         | Mạng nơ ron học sâu         |
| 4   | CNN         | Mạng nơ ron tích chập       |
| 5   | ASR         | Nhận dạng giọng nói tự động |

# CHƯƠNG 1: TỔNG QUAN VỀ ĐIỆN MẶT TRỜI

## 1.1 Giới thiệu

### 1.1.1 Khái niệm về điện mặt trời

Nguồn năng lượng tốt và lớn nhất có thể tận dụng được chính là Mặt trời. Để sử dụng năng lượng Mặt Trời không gây ảnh hưởng gì đến môi trường. Sử dụng năng lượng mặt trời không gây ra khí và nước độc hại, vì vậy không làm ô nhiễm môi trường và hiệu ứng nhà kính.

Cách để thu nhận và lưu trữ năng lượng Mặt Trời về cơ bản có hai phương pháp: Phương pháp chủ động sử dụng các thiết bị đặc biệt để thu bức xạ nhiệt và sử dụng các hệ thống quạt và máy bơm để phân phối nhiệt. Phương pháp thụ động sử dụng các nguyên tắc thu giữ nhiệt trong cấu trúc và vật liệu của các công trình xây dựng. Phương pháp thụ động có lịch sử phát triển lâu đời hơn, trong khi phương pháp chủ động chỉ mới được phát triển chủ yếu trong thế kỷ 20.

Các ứng dụng chính của năng lượng mặt trời là.

**Nhiệt Mặt Trời:** chuyển bức xạ Mặt Trời thành nhiệt năng, được ứng dụng trong hệ thống sưởi, hoặc để đun nước tạo hơi quay turbin điện.

**Điện Mặt Trời:** chuyển bức xạ Mặt Trời thành điện năng (hay còn gọi là quang điện-photovoltaics) .

### 1.1.2 Nguyên tắc cơ bản của bức xạ mặt trời

Nguồn năng lượng cho các quá trình phong hóa, bào mòn, vận chuyển, bồi tụ trên Trái Đất, cũng như chiếu sáng, sưởi ấm cho các hành tinh trong hệ Mặt Trời. Bức xạ mặt trời là dòng vật chất và năng lượng của Mặt Trời phát ra.

Bức xạ mặt trời gồm bức xạ hạt và bức xạ điện từ:

Bức xạ hạt hay còn gọi là gió mặt trời chủ yếu gồm các Bức xạ hạt hay còn gọi là gió Mặt Trời chủ yếu gồm các proton và electron. Đa phần thì chúng có hại cho các sinh vật, nhưng Trái Đất đã có tầng ozone bao phủ ngăn được phần nào ảnh hưởng có hại.

Năng lượng bức xạ hạt của Mặt Trời thường thấp hơn năng lượng bức xạ nhiệt 107 lần, và thâm nhập vào tầng khí quyển không quá 90 km. Khi đến gần Trái Đất, nó có vận tốc tới 300-1.525 km/s và mật độ 5-80 ion/cm<sup>3</sup>

Bức xạ điện từ có hai dạng: bức xạ trực tiếp và bức xạ khuếch tán. Có bước sóng khá rộng từ bức xạ gamma đến sóng vô tuyến với năng lượng cực đại ở vùng quang phổ khả kiến. Đây chính là nguồn năng lượng chủ yếu để chiếu sáng và duy trì các hoạt động sinh hóa trên Trái Đất. Khi qua khí quyển Trái Đất, các bức xạ sóng ngắn có hại cho sự sống gần như bị tầng ozone hấp thụ hoàn toàn. Ngày nay do công nghiệp phát triển, các chất CFC thải vào khí quyển đang huỷ hoại dần dần tầng ozone, tạo ra nguy cơ bức xạ sóng ngắn sẽ tiêu diệt sự sống trên Trái Đất.

Chùm tia truyền thẳng từ mặt trời gọi là bức xạ trực xạ. Tổng hợp các tia trực xạ và tán xạ gọi là tổng xạ. Mật độ dòng bức xạ trực xạ ở ngoài lớp khí quyển, tính đối với với  $1\text{m}^2$  bề mặt đặt vuông góc với tia bức xạ. Do khoảng cách giữa trái đất và mặt trời thay đổi theo mùa trong năm nên  $\beta$  cũng thay đổi do đó  $q$  cũng thay đổi nhưng độ thay đổi này không lớn lắm nên có thể xem  $q$  là không đổi và được gọi là hằng số mặt trời. Khi truyền qua lớp khí quyển bao bọc quanh trái đất các chùm tia bức xạ bị hấp thụ và tán xạ bởi tầng ôzôn, hơi nước và bụi trong khí quyển, chỉ một phần năng lượng được truyền trực tiếp tới trái đất. Đầu tiên ôxy phân tử bình thường  $\text{O}^2$  phân ly thành ôxy nguyên tử  $\text{O}$ , để phá vỡ liên kết phân tử đó, cần phải có các photon bước sóng ngắn hơn  $0,18\text{ }\mu\text{m}$ , do đó các photon (xem bức xạ như các hạt rời rạc - photon) có năng lượng như vậy bị hấp thụ hoàn toàn. Chỉ một phần các nguyên tử ôxy kết hợp thành các phân tử, còn đại đa số các nguyên tử tương tác với các phân tử ôxy khác để tạo thành phân tử ôzôn  $\text{O}^3$ . ôzôn cũng hấp thụ bức xạ tử ngoại nhưng với mức độ thấp hơn so với ôxy, dưới tác dụng của các photon với bước sóng ngắn hơn  $0,32\text{ }\mu\text{m}$ , sự phân tách  $\text{O}^3$  thành  $\text{O}^2$  và  $\text{O}$  xảy ra. Như vậy hầu như toàn bộ năng lượng của bức xạ tử ngoại được sử dụng để duy trì quá trình phân ly và hợp nhất của  $\text{O}$ ,  $\text{O}^2$  và  $\text{O}^3$ , đó là một quá trình ổn định. Do quá trình này, khi đi qua khí quyển, bức xạ tử ngoại biến đổi thành bức xạ với năng lượng nhỏ hơn. Các bức xạ với bước sóng ứng với các vùng nhìn thấy và vùng hồng ngoại của phổ tương tác với các phân tử khí và các hạt bụi của không khí nhưng không phá vỡ các liên kết của chúng, khi đó các photon bị tán xạ khá đều theo mọi hướng và một số photon quay trở lại không gian vũ trụ. Bức xạ chịu dạng tán xạ đó chủ yếu là bức xạ có bước sóng ngắn nhất. Sau khi phản xạ từ các phần khác nhau của khí quyển bức xạ tán xạ đi đến chúng ta mang theo màu xanh lam của bầu trời trong sáng và



có thể quan sát được ở những độ cao không lớn. Các giọt nước cũng tán xạ rất mạnh bức xạ mặt trời.

Bức xạ mặt trời khi đi qua khí quyển còn gặp một trở ngại đáng kể nữa đó là do sự hấp thụ của các phân tử hơi nước, khí cacbonic và các hợp chất khác, mức độ của sự hấp thụ này phụ thuộc vào bước sóng, mạnh nhất ở khoảng giữa vùng hồng ngoại của phổ. Phần năng lượng bức xạ mặt trời truyền tới bề mặt trái đất trong những ngày quang đăng (không có mây) ở thời điểm cao nhất vào khoảng  $1000\text{W/m}^2$ . Yếu tố cơ bản xác định cường độ của bức xạ mặt trời ở một điểm nào đó trên trái đất là quãng đường nó đi qua. Sự mất mát năng lượng trên quãng đường đó gắn liền với sự tán xạ, hấp thụ bức xạ và phụ thuộc vào thời gian trong ngày, mùa, vị trí địa lý. Các mùa hình thành là do sự nghiêng của trục trái đất đối với mặt phẳng quỹ đạo của nó quanh mặt trời gây ra. Góc nghiêng vào khoảng  $66,5^\circ$  và thực tế xem như không đổi trong không gian. Sự định hướng như vậy của trục quay trái đất trong chuyển động của nó đối với mặt trời gây ra những sự dao động quan trọng về độ dài ngày và đêm trong năm. Công suất này bằng 4.1013 lần tổng công suất điện toàn thế giới hiện nay, vào khoảng  $P = 1013\text{W}^2$ . Nguyên tắc tận thu bức xạ mặt trời Cường độ bức xạ trên mặt đất chủ yếu phụ thuộc 2 yếu tố: góc nghiêng của các tia sáng đối với mặt phẳng bề mặt đã cho và độ dài đường đi của các tia sáng trong khí quyển nói chung là phụ thuộc vào độ cao của mặt trời (góc giữa phương từ điểm quan sát đến mặt trời và mặt phẳng nằm ngang đi qua điểm đó). Yếu tố cơ bản xác định cường độ của bức xạ mặt trời ở một điểm nào đó trên trái đất là quãng đường nó đi qua. Sự mất mát năng lượng trên quãng đường đó gắn liền với sự tán xạ, hấp thụ bức xạ và phụ thuộc vào thời gian trong ngày, mùa, vị trí địa lý. Quan hệ giữa bức xạ mặt trời ngoài khí quyển và thời gian trong năm có thể xác định theo phương trình. Tính toán góc tới của bức xạ trực xạ Hệ số khối không khí:  $m$ , là tỷ số giữa khối lượng khí quyển theo phương tia bức xạ truyền qua và khối lượng khí quyển theo phương thẳng đứng (tức là khi mặt trời ở đỉnh). Như vậy  $m = 1$  khi mặt trời ở đỉnh,  $m = 2$  khi góc đỉnh  $\theta_z$  là  $60^\circ$ . Đối với các góc đỉnh từ  $0 - 70^\circ$  có thể xác định gần đúng  $m = 1/\cos\theta_z$ . Còn đối với các góc  $\theta_z$  thì độ cong của bề mặt trái đất phải được đưa vào tính toán. Riêng đối với trường hợp tính toán bức xạ mặt trời ngoài khí quyển  $m = 0$ .

Trực xạ: là bức xạ mặt trời nhận được khi không bị bầu khí quyển phát tán. Đây là dòng bức xạ có hướng và có thể thu được ở các bộ thu kiểu tập trung (hội tụ).

Tán xạ: là bức xạ mặt trời nhận được sau khi hướng của nó đã bị thay đổi do sự phát tán của bầu khí quyển (trong một số tài liệu khí tượng, tán xạ còn được gọi là bức xạ của bầu trời, ở đây cần phân biệt tán xạ của mặt trời với bức xạ hồng ngoại của bầu khí quyển phát ra).

Tổng xạ: là tổng của trực xạ và tán xạ trên một bề mặt (phổ biến nhất là tổng xạ trên một bề mặt nằm ngang, thường gọi là bức xạ cầu trên bề mặt). Cường độ bức xạ ( $W/m^2$ ): là cường độ năng lượng bức xạ mặt trời đến một bề mặt tương ứng với một đơn vị diện tích của bề mặt. Cường độ bức xạ cũng bao gồm cường độ bức xạ trực xạ  $E_{trx}$ , cường độ bức xạ tán xạ  $E_{tx}$  và cường độ bức xạ quang phổ  $E_{qp}$ . Năng lượng bức xạ ( $J/m^2$ ): là năng lượng bức xạ mặt trời truyền tới một đơn vị diện tích bề mặt trong một khoảng thời gian, như vậy năng lượng bức xạ là một đại lượng bằng tích phân của cường độ bức xạ trong một khoảng thời gian nhất định (thường là 1 giờ hay 1 ngày). Giờ mặt trời: là thời gian dựa trên chuyển động biểu kiến của mặt trời trên bầu trời, với quy ước giờ mặt trời chính Ngọ là thời điểm mặt trời đi qua thiên đỉnh của người quan sát. Giờ mặt trời là thời gian được sử dụng trong mọi quan hệ về góc mặt trời, nó không đồng nghĩa với giờ theo đồng hồ. Quan hệ hình học giữa một mặt phẳng bố trí bất kỳ trên mặt đất và bức xạ của mặt trời truyền tới, tức là vị trí của mặt trời so với mặt phẳng đó có thể được xác định theo các góc đặc trưng:

Góc vĩ độ  $\varphi$ : vị trí góc tương ứng với vĩ độ về phía bắc hoặc về phía nam đường xích đạo trái đất, với hướng phía bắc là hướng dương.  $-90^\circ \leq \varphi \leq 90^\circ$

Góc nghiêng  $\beta$ : góc giữa mặt phẳng của bề mặt tính toán và phương nằm ngang  $0^\circ \leq \beta \leq 180^\circ$  ( $\beta > 90^\circ$  nghĩa là bề mặt nhận bức xạ hướng xuống phía dưới).

Góc phương vị của bề mặt  $\gamma$ : góc lệch của hình chiếu pháp tuyến bề mặt trên mặt phẳng nằm ngang so với đường kinh tuyến. Góc  $\gamma = 0$  nếu bề mặt quay về hướng chính nam,  $\gamma$  lấy dấu (+) nếu bề mặt quay về phía tây và lấy dấu (-) nếu bề mặt quay về phía đông.  $-180^\circ \leq \gamma \leq 180^\circ$

## 1.2 Một số hệ thống điện mặt trời phổ biến

**Hệ thống điện mặt trời độc lập:** Đây là hệ thống phổ biến nhất hiện nay ở Việt Nam vì đáp ứng được các nhu cầu sử dụng điện tại các khu vực chưa có điện, điện chập chờn, khi mất điện hoặc đơn giản là sử dụng điện mặt trời cho một số ít các thiết bị trong gia đình. Là hệ thống chỉ dùng điện lấy được từ năng lượng mặt trời cung cấp điện cho các thiết bị điện độc lập sử dụng cả ngày lẫn đêm. Bao gồm các thành phần sau: pin mặt trời, thiết bị điều khiển nạp, ắc quy lưu trữ, có hoặc không có bộ đổi điện DC – AC.

**Hệ thống điện mặt trời độc lập có bù lưới:** là hệ thống khi yếu điện, đảm bảo các thiết bị đã đấu nối vào hệ thống điện mặt trời luôn đủ điện dùng dù trời không nắng hoặc nhu cầu sử dụng tăng cao. Điện mặt trời độc lập nhưng có kết hợp bộ bù lưới thông minh để xử lý.

**Hệ thống điện mặt trời độc lập kết hợp hòa lưới:** là hệ thống có kết hợp hòa lưới khi hệ thống ắc quy dự trữ đã được nạp đầy. Khi sử dụng không hết nguồn điện mặt trời và không còn nạp thêm cho ắc quy được nữa thì hệ thống tự động chuyển sang hòa phần điện mặt trời đó vào lưới. Khi lượng điện trong ắc quy giảm xuống, hệ thống sẽ quay lại sạc đầy ắc quy.

**Hệ thống điện mặt trời trực tiếp:** là hệ thống chỉ bao gồm các tấm pin mặt trời và bộ điều khiển sạc, và lấy điện DC từ bộ điều khiển sạc cấp cho tải. Khi không có mặt trời hệ thống sẽ tự ngắt để bảo vệ quá tải.

**Hệ thống điện mặt trời nối lưới:** Là hệ thống sử dụng song song với lưới điện, lượng điện phát từ pin mặt trời khi không sử dụng sẽ được chuyển đổi sang hòa vào lưới điện quốc gia, phù hợp với mô hình sử dụng công tơ 2 chiều để bán lại điện khi không sử dụng.

**Hệ thống điện mặt trời nối lưới có dự trữ:** là hệ thống vận hành song song với lưới điện và có hệ thống lưu điện khi không sử dụng hết nguồn phát từ pin mặt trời.

**Hệ thống điện mặt trời độc lập kết hợp:** là hệ thống kết hợp với một số nguồn phát khác như tua bin gió hoặc máy phát điện dùng nhiên liệu. Ưu điểm hệ thống này là đảm bảo nguồn phát điện liên tục hơn chỉ sử dụng hệ thống điện mặt trời, nhược điểm là phức tạp và giá cao, chi phí cho bảo trì cũng lớn.

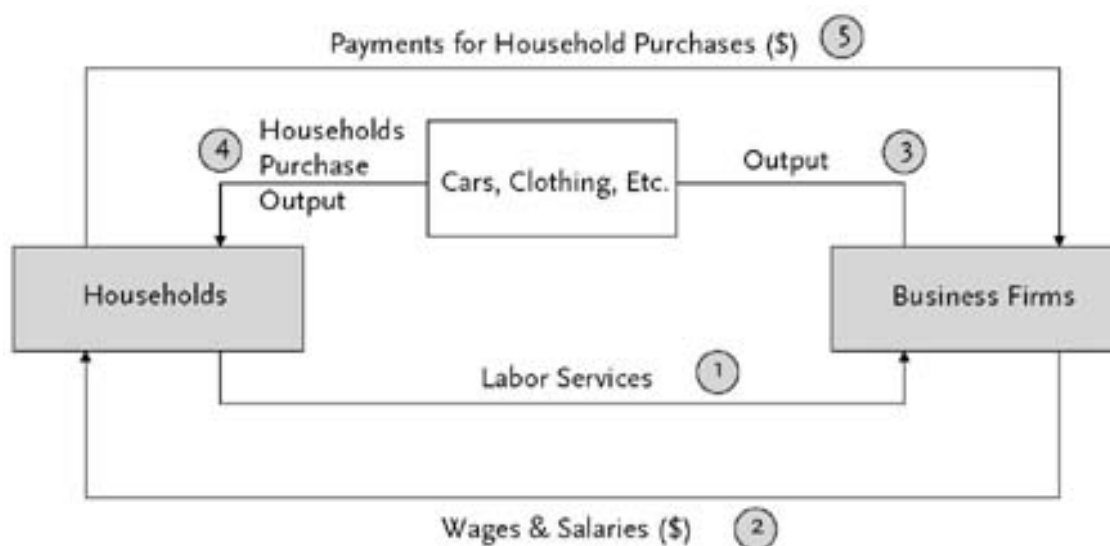
**Hệ thống điện mặt trời độc lập thông minh:** là hệ thống kết hợp thông minh nhiều chế độ chạy giữa năng lượng mặt trời, năng lượng gió, điện lưới và máy phát điện dự phòng (nếu có). Có thể tự lập trình chọn chế độ chạy theo nhu cầu. Hệ thống đảm bảo nguồn điện thông minh 24/24 với bất kỳ lý do nào.

### 1.3 Một số phương pháp dự báo

#### 1.3.1 Phương pháp thống kê kinh tế

##### 1.3.1.1 Dự báo kinh tế sử dụng mô hình kinh tế lượng

Mô hình kinh tế lượng là một trong những công cụ kinh tế sử dụng để dự báo sự phát triển tương lai của nền kinh tế. Nói một cách đơn giản, kinh tế thông thường đo lường mối quan hệ giữa các biến trong quá khứ như chi tiêu của người tiêu dùng, thu nhập hộ gia đình, mức thuế suất, lãi suất, việc làm, và sau đó dự báo những thay đổi trong một số biến sẽ ảnh hưởng đến tương lai.



**Hình 1.1: Mô hình lương và chi tiêu trong gia đình.**

Hình 1.1 trên mô tả các chi phí trong gia đình liên quan và ảnh hưởng tới thu nhập. Ở bảng 1.1 dưới đây mô tả sự thay đổi giữa thu nhập và tiêu dùng trong gia đình về đồ dùng cá nhân như giày dép, quần áo...

**Bảng 1.1: Tỷ lệ phần trăm thu nhập và tiêu dùng thời trang**

| Year    | % of Household Income |             |
|---------|-----------------------|-------------|
|         | Total                 | Total - 100 |
| 1995    | 4.5                   | 2.6         |
| 1996    | 4.4                   | 2.6         |
| 1997    | 4.3                   | 2.6         |
| 1998    | 4.2                   | 2.7         |
| 1999    | 4.3                   | 2.8         |
| 2000    | 4.1                   | 2.7         |
| 2001    | 4.0                   | 2.6         |
| 2002    | 3.9                   | 2.6         |
| Average | 4.21                  | 2.65        |

Từ mối quan hệ giữa những dữ liệu trên cho thấy, người Mỹ chi tiêu nhiều hơn vào quần áo và giày thu nhập hộ gia đình của họ cao hơn, nhưng họ dành ít nhất 100 tỷ \$ mỗi năm. Và dự báo khá chính xác khi tổng số đó sẽ được chi tiêu là: 100 \$ tỷ cộng thêm 2,65 phần trăm thu nhập hộ gia đình. Ở dạng phương trình, điều này được thể hiện bằng  $C = 100 + 0.0265W$ . Sử dụng dữ liệu để xác định hoặc ước tính tất cả các giá trị tham số trong mô hình là bước quan trọng có thể biến các mô hình kinh tế toán học thành một mô hình kinh tế lượng.

Kinh tế lượng là việc áp dụng các phương pháp thống kê số liệu kinh tế và được mô tả như những nhánh của kinh tế nhằm đưa ra nội dung thực nghiệm để quan hệ kinh tế. Chính xác hơn, đó là "sự phân tích định lượng của các hiện tượng kinh tế thực tế dựa trên sự phát triển đồng thời của các lý thuyết và quan sát, liên quan bằng các phương pháp thích hợp của suy luận. Kinh tế lượng ứng dụng sử dụng kinh tế lượng lý thuyết và số liệu thực tế để đánh giá các lý thuyết kinh tế, phát triển mô hình kinh tế, phân tích lịch sử kinh tế, và dự báo".

Dự báo từ một mô hình tuyến tính đơn giản có thể dễ dàng thu được bằng cách sử dụng phương trình

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad (1-1)$$

Trong đó  $x$  là biến độc lập và  $y$  là biến phụ thuộc. Hay nói cách khác,  $x$  là các yếu tố liên quan tới dự báo,  $y$  là giá trị của dự báo tương ứng với các nhân tố  $x$  [9].

### 1.3.1.2 Mô hình hồi quy

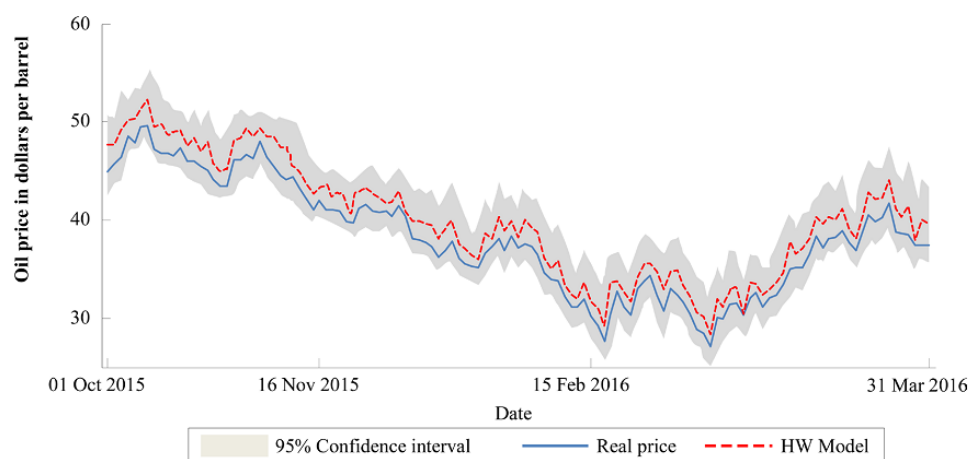
Z. Ismail và các cộng sự đã sử dụng mô hình hồi quy bội (multiple linear regression) để dự báo giá vàng tại Malaysia. Trong nghiên cứu này, họ sử dụng các yếu tố kinh tế như lạm phát, biến động giá tiền tệ và những người khác. Sau sự giảm giá của đồng đô la Mỹ, nhà đầu tư đổ tiền vào vàng bởi vì vàng đóng một vai trò quan trọng như một mục đầu tư ổn định. Do sự gia tăng nhu cầu đối với vàng trong Malaysia và các quốc gia khác trên thế giới. Họ cũng nhận thấy rằng có thể phát triển một mô hình phản ánh cấu trúc và của thị trường vàng và chuyển động dự báo giá vàng. Các cách tiếp cận phù hợp nhất là mô hình hồi quy bội MLR để nghiên cứu về mối quan hệ giữa một biến duy nhất phụ thuộc và một hoặc nhiều hơn một các biến độc lập, như trường hợp này với giá vàng là biến phụ thuộc duy nhất. Một mô hình đơn giản gọi là "dự báo-1" đã được coi là một mô hình chuẩn để đánh giá hiệu quả của mô hình

Tetyana Kuzhda đã đề xuất mô hình dự báo bán lẻ dựa trên mô hình hồi quy bội. Họ cho rằng, môi trường bán lẻ gần đây có sự thay đổi, sự thay đổi này được gây ra bởi chi phí thu nhập và quảng cáo. Mô hình của họ được xây dựng dựa trên các tác nhân ảnh hưởng (stage) bao gồm: ảnh hưởng kinh tế mức vĩ mô (lạm phát, cung cầu, tỉ giá hối đoái, chi tiêu Chính phủ, xuất nhập khẩu,...), yếu tố giải thích thay đổi trong kinh tế xã hội (doanh số bán lẻ, chi phí quảng cáo,...). Từ đó họ xây dựng mô hình bằng cách thu thập những dữ liệu của những tác nhân trên và xây dựng hàm toán học giải quyết vấn đề dự báo bán lẻ.

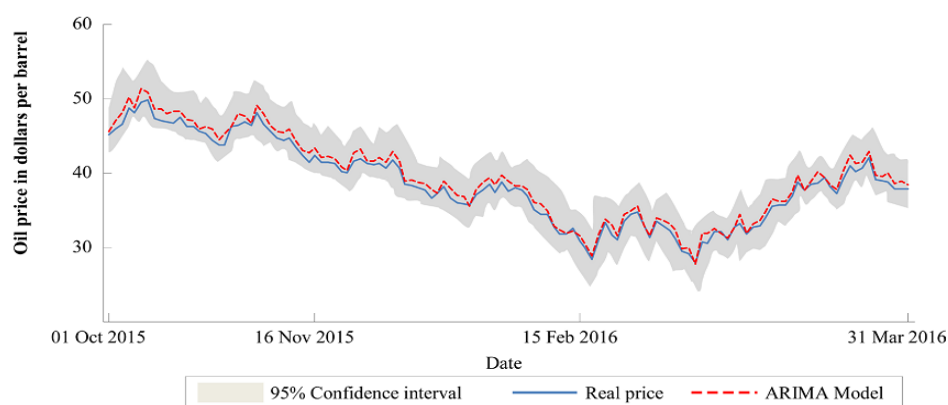
Vincenzo Bianco và các cộng sự dự báo về mức tiêu thụ điện tại Italy sử dụng mô hình hồi quy tuyến tính. Họ cho rằng những ảnh hưởng các biến số kinh tế và nhân khẩu học về tiêu thụ điện hàng năm ở Ý có thể là nhân tố tác động lên mức tiêu thụ điện. Họ đưa ra khoảng thời gian xem xét các dữ liệu lịch sử được từ năm 1970 đến năm 2007. Mô hình hồi quy được xây dựng dựa trên các tham số: lịch sử sử dụng tiêu thụ điện, tổng sản phẩm trong nước (GDP), tổng thu nhập bình quân đầu người (GDP bình quân đầu người) và một phần dân số. Trong phần thực nghiệm, họ có so sánh đánh giá với dự báo quốc gia dựa trên mô hình phức tạp hơn cho thấy tỉ lệ sai số là nhỏ với độ lệch chỉ  $\pm 1\%$ .

### 1.3.1.3 Mô hình chuỗi thời gian

Gurudeo Anand Tularam đã đề xuất phương pháp dự báo giá dầu dựa trên mô hình chuỗi thời gian, họ chứng minh rằng những dữ liệu về giá xăng dầu trong quá khứ ảnh hưởng tới giá dầu trong tương lai. Trong bài báo của họ cũng thực hiện việc đánh giá độ chính xác của mô hình với giá dầu thực. Ba loại mô hình đơn biến được sử dụng: làm mịn số mũ (ES), Holt – Winters (HW), tự hồi quy trung bình trượt (ARIMA). Kết quả dự báo với độ chính xác 95% khi dự báo giá dầu thô West Texas Intermediate.



**Hình 1.2: Kết quả thử nghiệm với mô hình HW**



**Hình 1.3: Kết quả thử nghiệm với mô hình ARIMA**

Wen-Hua Cui và các cộng sự trong năm 2014 đã thực hiện công việc dự báo về dòng tiền trong ngân hàng nhằm mục đích tăng cường thông tin giữa dòng tiền kinh doanh và quỹ. Đầu tiên họ dùng phương pháp trung bình trượt, tiếp theo làm

mìn số mũ để dự đoán. Hai phương pháp này đều kết hợp với chuỗi thời gian trong dự báo dòng tiền. Kết quả thực nghiệm được thực hiện với dòng tiền thực trong khoảng thời gian quý 1 năm 2011 và 2012 cho thấy rằng, phương pháp của họ đề xuất có độ chính xác khá cao.

#### 1.4 Phương pháp đánh giá dự báo

Đánh giá dự báo là đánh giá độ chính xác của dự báo. Do đó, cần xác định được các lỗi khi so sánh giữa giá trị thực và kết quả cho ra bởi mô hình.

Lỗi phụ thuộc độ đo (scale – dependent error)

Các lỗi dự báo được tính:

$$e_i = y_i - \hat{y}_i \quad (1-2)$$

Trong đó dựa trên cùng một độ đo, cùng một dữ liệu. Hai phương pháp phụ thuộc và thông dụng nhất được dựa trên các lỗi tuyệt đối hoặc lỗi bình phương:

$$\text{Mean absolute error: MAE} = \text{mean}(|e_i|) \quad (1-3)$$

$$\text{Root mean squared error: RMSE} = \sqrt{\text{mean}(e_i^2)}. \quad (1-4)$$

Khi so sánh các phương pháp dự báo trên một tập dữ liệu duy nhất, MAE được dùng phổ biến vì dễ dàng trong tính toán.

Lỗi tỉ lệ (percentage error)

$$\text{Các lỗi tỷ lệ được tính bởi: } p_i = 100e_i / y_i \quad (1-5)$$

Lỗi tỉ lệ thuận lợi hơn lỗi phụ thuộc độ đo. Vì vậy thường được sử dụng để so sánh hiệu suất dự báo bộ dữ liệu khác nhau. Các biện pháp thường được sử dụng nhất là:

$$\text{Mean absolute percentage error: MAPE} = \text{mean}(|p_i|) \quad (1-6)$$

Lỗi độ đo:

Lỗi độ đo được đề xuất bởi Hyndman và Koehler (2006) như một sự thay thế cho việc sử dụng các lỗi tỷ lệ phần trăm khi so sánh chính xác dự báo trên loạt trên quy mô khác nhau. Họ đề xuất nhân rộng các lỗi dựa trên đạo tạo MAE từ một phương pháp dự báo đơn giản. Đối với một chuỗi thời gian không theo mùa, một cách hữu ích để xác định một lỗi có quy mô sử dụng dự báo Naive:



$$q_j = \frac{e_j}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \quad (1-7)$$

Bởi vì tử số và mẫu số đều liên quan đến các giá trị trên độ đo của dữ liệu gốc,  $q_j$  là độc lập với quy mô của dữ liệu. Một lỗi có độ đo nhỏ hơn một nếu nó xuất phát từ một dự báo tốt hơn so với dự báo Naive trung bình tính trên dữ liệu huấn luyện. Ngược lại, nó lớn hơn một nếu theo dự báo là kém hơn so với dự báo naive trung bình tính trên dữ liệu huấn luyện. Đối với chuỗi thời gian theo mùa, một lỗi có thể được xác định bằng:

$$q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|} \quad (1-8)$$

Đối với dữ liệu chéo:

$$q_j = \frac{e_j}{\frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}|} \quad (1-9)$$

Trong trường hợp này, việc so sánh là với dự báo trung bình. Các sai số trung bình có quy mô tuyệt đối là:

$$\text{MASE} = \text{mean}(|q_j|) \quad (1-10)$$

## 1.5 Kết luận chương 1

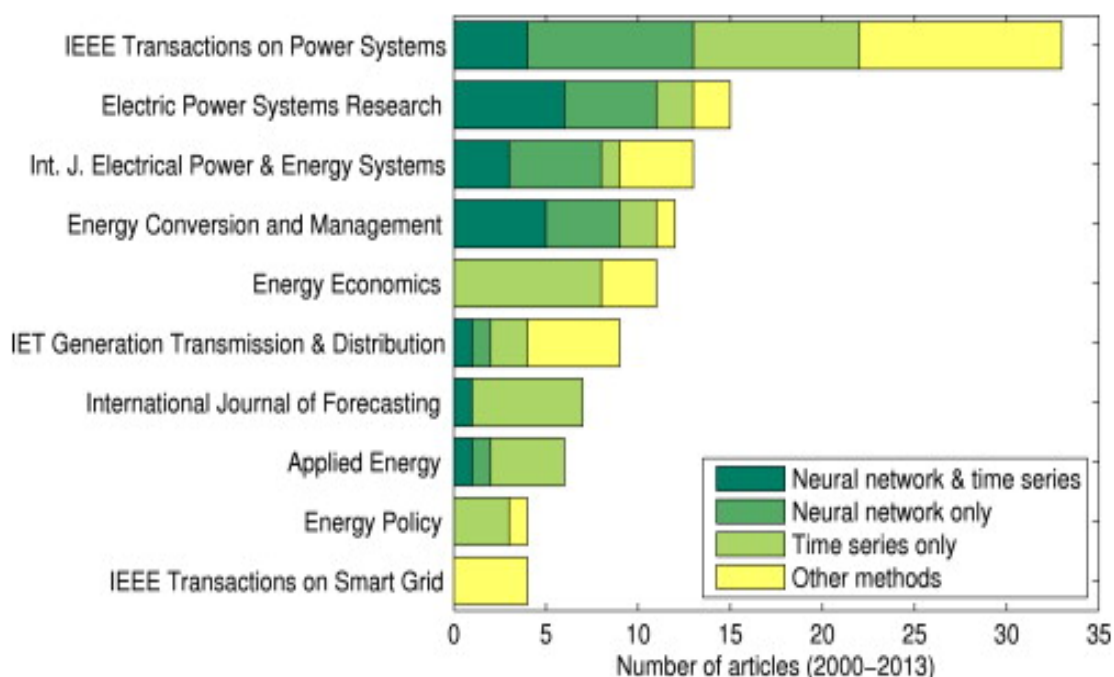
Trong chương này, luận văn đã trình bày tổng quan các khái niệm cơ bản liên quan tới điện mặt trời, các phương pháp dự báo, dự báo kinh tế. Các phương pháp thường sử dụng trong dự báo kinh tế bao gồm các phương pháp liên quan tới kinh tế lượng và học máy. Trong luận văn sử dụng các phương pháp định lượng để dự báo, do đó lý thuyết về phương pháp định lượng được trình bày nhiều hơn. Cuối chương, luận văn cũng đưa ra một số phương pháp xác định lỗi dự báo nhằm mục đích đánh giá độ chính xác của các dự báo dựa trên sai lệch kết quả giữa mô hình và giá trị thực tế của dự báo.

## CHƯƠNG 2: DỰ BÁO NHU CẦU SỬ DỤNG ĐIỆN MẶT TRỜI

### SỬ DỤNG PHƯƠNG PHÁP HỌC MÁY

#### 2.1 Một số phương pháp liên quan

Nhận thức được sự quan trọng của dự báo nhu cầu sử dụng điện, các phương pháp đề xuất cho dự báo nhu cầu tăng nhanh trong những năm gần đây. Dự báo đến năm 2020, nhu cầu sử dụng điện của Việt Nam sẽ tăng 3,5 lần so với hiện nay - đó là nhận định của ông Kimio Yoshida, Trưởng nhóm nghiên cứu tổng sơ đồ sử dụng năng lượng hiệu quả tại Việt Nam của Tổ chức Hợp tác Quốc tế Nhật Bản (JICA). Hình dưới đây mô tả một số các phương pháp được sử dụng nhiều trong dự báo nhu cầu sử dụng điện. (thống kê từ cơ sở dữ liệu một số tạp chí)



**Hình 2.1: Một số phương pháp tiếp cận dự báo**

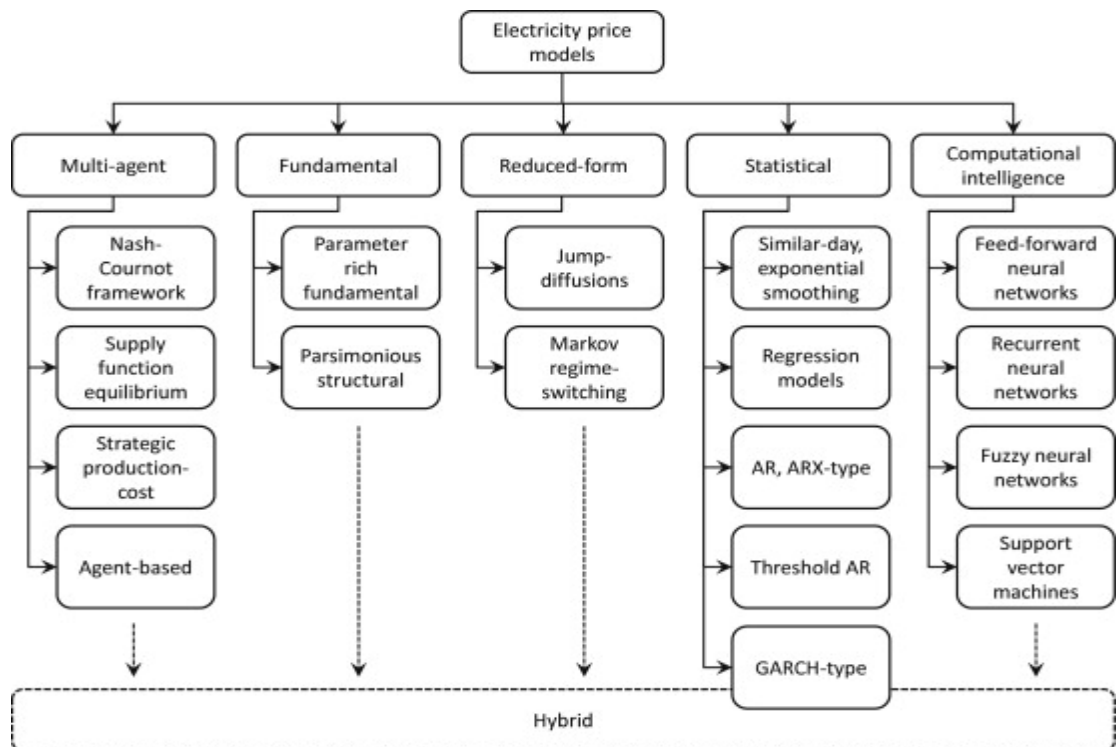
Vào cuối những năm 1990, các phương pháp tính toán tin học và các mạng nơ ron là một chủ đề nóng cho các nhà nghiên cứu, các học giả quan tâm phát triển các thuật toán. Một số công bố điển hình trong những năm 90 bao gồm:

- Shahidehpour et al. (Năm 2002) thảo luận về các vấn đề cơ bản của việc định giá và dự báo (biến nội sinh, các biến ngoại sinh), mô tả một mô đun dự báo dựa trên mạng nơ-ron và bình luận về đánh giá hiệu năng.
- Weron (2006) cung cấp tổng quan các phương pháp tiếp cận mô hình hoá, sau đó tập trung vào các ứng dụng thực tiễn của các phương pháp thống kê

cho dự báo (ARMA, ARMAX,) dự báo khoảng thời gian, và chuyển sang các mô hình định lượng để định giá phái sinh.

- Zareipour (2008) bắt đầu bằng việc xem xét các mô hình chuỗi thời gian tuyến tính (ARIMA, ARX, ARMAX) và các mô hình phi tuyến (hồi quy splines, mạng nơ-ron), sau đó sử dụng chúng để dự báo hàng giờ trên thị trường điện Ontario.

Tiếp theo đó, có khá nhiều các phương pháp được sử dụng để dự báo, có thể phân thành các nhóm phương pháp như sau:



**Hình 2.2: Phân nhóm các phương pháp dự báo nhu cầu sử dụng điện**

### 2.1.1 Multi-agent models

Dự báo dài hạn nhu cầu sử dụng điện được coi như là việc đơn giản nhưng tốt kém, nó thường liên quan tới dự báo trung bình và dài hạn và kết hợp các ước tính về nhu cầu cung cấp. Các mô hình chi phí (mô hình chi phí sản xuất) có khả năng dự báo nhu cầu sử dụng theo giờ. Tuy nhiên, cách thức này lại bỏ qua những vấn đề phức tạp của thị trường. Các mô hình này thường coi thị trường là khá ổn định, không bị chi phối bởi các chiến lược hay sự cạnh tranh bất thường khác. Các mô hình này có thể dự đoán khá chính xác mức nhu cầu dự kiến của một thị trường điện với tính cân bằng cao.

Ventosa và các cộng sự (2005) xác định ba xu hướng mô hình thị trường điện chính: tối ưu hóa, cân bằng và các mô hình mô phỏng. Trong phân loại của họ, các mô hình tối ưu hóa tập trung vào vấn đề tối đa hóa lợi nhuận cho một trong những công ty cạnh tranh trên thị trường. Họ đã sử dụng mô hình đa tác tử mô phỏng cho các công ty cạnh tranh để quyết định được nhu cầu của thị trường dựa trên những sự tương tác trên những agent.

Trong Nash – Cournot framework, điện được coi là một sản phẩm đồng nhất và cân bằng thị trường được xác định thông qua các quyết định về khả năng của nhà cung cấp. Tuy nhiên, các mô hình này có xu hướng cung cấp nhu cầu sử dụng cao hơn giá trị được quan sát trong thực tế. Các nhà nghiên cứu đã giải quyết vấn đề này bằng cách giới thiệu khái niệm các biến phán đoán nhằm mục đích đại diện cho thực tế là các đối thủ cạnh tranh sẽ sản xuất điện nhiều hơn khi thấy nhu cầu sử dụng điện cao hơn. Ruibal và Mazumdar (2008) cung cấp một trong số rất ít ứng dụng của framework này cho EPF. Một mô hình ngẫu nhiên dựa trên cơ sở nhu cầu sử dụng thâu được đề xuất để dự đoán nhu cầu sử dụng điện theo giờ và nhu cầu sử dụng trung bình trong một khoảng thời gian nhất định. Hai nguồn không chắc chắn được xem xét: sự sẵn có của các đơn vị phát điện và nhu cầu. Kết quả cho thấy khi số lượng các doanh nghiệp trên thị trường giảm, giá trị dự kiến của giá sẽ tăng lên đáng kể. Sự khác nhau của mô hình Cournot cũng tăng lên Ruibal và Mazumdar cũng chứng minh rằng một dự báo nhiệt độ chính xác có thể làm giảm đáng kể dự đoán của dự báo nhu cầu sử dụng điện.

### **2.1.2 Các mô hình cơ bản**

Coulon và Howison (2009) phát triển một mô hình cơ bản cho nhu cầu sử dụng điện tại chỗ, dựa trên các quy trình ngẫu nhiên cho các yếu tố cơ bản (giá nhiên liệu, nhu cầu năng lượng và khả năng sẵn có năng lượng), cũng như một hình thức tham số cho trình điều khiển để giá điện. Sử dụng dữ liệu giá thâu, họ tìm thấy mối tương quan giữa giá thâu và giá nhiên liệu tương ứng. Sử dụng mô hình ngẫu nhiên của chào giá, Carmona, Coulon và Schwarz (2013) đã biến nhu cầu về điện và giá nhiên liệu để dự báo giá điện. Trong một bối cảnh tương tự, Aïd, Campi và Langrené (2013) đã xây dựng một mô hình cấu trúc rủi ro trong đó có một chức năng khan hiếm được đưa ra để cho phép lệch giá tại chỗ từ giá nhiên liệu cận biên,

dẫn đến tăng giá. Giống như Carmona và các cộng sự (2013), họ tập trung vào việc định giá và bảo hiểm rủi ro ảnh hưởng tới giá điện.

### **2.1.3 Mô hình reduced-form**

Một tính năng phổ biến của mô hình reduced – form là không cung cấp dự báo giá mỗi giờ chính xác mà thường là dự báo sử dụng phân phối cận biên tại điểm thời gian trong tương lai. Trong mô hình này, họ sử dụng các luật tính toán (Härdle & Weron, 2011) hoặc các cơ chế về kinh tế lượng (đột ngột thay đổi giá, xem Hamilton, 2008). Theo một cách nào đó, các mô hình này được xem xét theo các mô hình chuyển đổi chế độ Markov.

### **2.1.4 Mô hình thống kê và kinh tế lượng**

Một trong các ứng dụng chủ yếu của mô hình kinh tế lượng là để dự báo hay dự đoán. Có hai nhóm phương pháp dự báo: Kinh tế lượng và chuỗi thời gian. Dự báo kinh tế lượng đặt cơ sở trên mô hình hồi quy để nối kết một (hay nhiều) biến phụ thuộc với một số biến độc lập. Dự báo chuỗi thời gian nối kết biến phụ thuộc với các giá trị của nó trong quá khứ và cố gắng sử dụng mối quan hệ này để dự báo biến phụ thuộc.

Một môi trường dự báo bao gồm ba thời đoạn. Một người khảo sát sử dụng một mẫu các quan sát và ước lượng mô hình. Các giá trị dự báo của biến phụ thuộc trong thời kỳ trong mẫu này còn được gọi là các giá trị thích hợp hóa. Các giá trị dự báo ngoài mẫu có thể là kiểm định hay tiên nghiệm. Dự báo kiểm định là đối với giai đoạn mà trong đó các giá trị thực tế của biến phụ thuộc và độc lập đều đã biết. Các giá trị dự báo như thế thường được so sánh với giá trị thực tế để đánh giá năng lực dự báo của mô hình. Các dự báo tiên nghiệm là dự báo cho tương lai với các giá trị của biến độc lập được dự báo từ các mô hình khác.

Dự báo có thể là có điều kiện hoặc không điều kiện. Khi các giá trị của các biến độc lập là biết trước thì ta có dự báo có điều kiện. Dự báo không điều kiện được tạo ra khi các giá trị của các biến ngoại sinh không được biết trước mà là được tạo ra từ bản thân mô hình hay từ một mô hình hỗ trợ khác.

Nhiều nghiên cứu trước đây đã sử dụng một mô hình ARIMA dựa trên chuyển đổi wavelet để dự báo giá điện. Dữ liệu lịch sử lần đầu tiên được phân chia

bằng cách sử dụng chuyển đổi wavelet trước khi áp dụng mô hình ARIMA. Sau đó, các kết quả dự báo thu được bằng cách sử dụng phép biến đổi wavelet nghịch đảo.

Mô hình ARIMA có tên tiếng Anh là Autoregressive Integrated Moving Average. Mô hình này lần đầu tiên được đưa ra bởi Box & Jenkins (1970). ARIMA được kết hợp bởi 3 thành phần chính: AR (thành phần tự hồi quy), I (tính dừng của chuỗi thời gian) và MA (thành phần trung bình trượt). Theo Gujarati (2004), để ước lượng mô hình ARIMA ta cần đi qua 4 bước chính sau:

- **Bước 1: Nhận dạng mô hình**

Mô hình ARIMA chỉ được áp dụng đối với chuỗi dừng. Mô hình có thể trình bày theo dạng AR, MA hay ARMA. Phương pháp xác định mô hình thường được thực hiện qua nghiên cứu chiều hướng biến đổi của hàm tự tương quan ACF hay hàm tự tương quan từng phần PACF.

**Chuỗi ARIMA không dừng:** cần phải được chuyển đổi thành chuỗi dừng trước khi tính ước lượng tham số bình phương tối thiểu. Việc chuyển đổi này được thực hiện bằng cách tính sai phân giữa các giá trị quan sát dựa vào giả định các phần khác nhau của các chuỗi thời gian đều được xem xét tương tự, ngoại trừ các khác biệt ở giá trị trung bình. Nếu việc chuyển đổi này không thành công, sẽ áp dụng tiếp các kiểu chuyển đổi khác (chuyển đổi logarithm chẳng hạn).

- **Bước 2: Ước lượng các tham số và lựa chọn mô hình**

Tính những ước lượng khởi đầu cho các tham số  $a_0, a_1, \dots, a_p, b_1, \dots, b_q$  của mô hình dự định. Sau đó xây dựng những ước lượng sau cùng bằng một quá trình lặp.

- **Bước 3: Kiểm định mô hình**

Sau khi các tham số của mô hình tổng quát đã xây dựng, ta kiểm tra mức độ chính xác và phù hợp của mô hình với dữ liệu. Chúng ta kiểm định phần dư  $(Y_t - \hat{Y}_t)$  và có ý nghĩa cũng như mối quan hệ các tham số. Nếu bất cứ kiểm định nào không thỏa mãn, mô hình sẽ nhận dạng lại các bước trên được thực hiện lại.

- **Bước 4: Dự báo**

Khi mô hình thích hợp với dữ liệu đã tìm được, ta sẽ thực hiện dự báo tại thời điểm tiếp theo  $t$ . Do đó, mô hình ARMA(p,q):

$$y(t+1) = a_0 + a_1 y(t) + \dots + a_p y(t-p+1) + e(t+1) + b_1 e(t) + \dots + b_q e(t-q+1)$$

## 2.1.5 Tính toán thông minh

### 2.1.5.1 Mạng Nơ-ron nhân tạo

ANN hứa hẹn đưa ra những kết quả rất chính xác cho thấy ANN đã được đào tạo với một bộ các tính năng đầu vào chính xác và đủ số liệu đầu vào.

Có lẽ lợi thế lớn nhất của ANNs là khả năng được sử dụng như một cơ chế xấp xỉ hàm tùy ý mà 'học' được từ các dữ liệu quan sát. Tuy nhiên, sử dụng chúng không đơn giản như vậy, và một sự hiểu biết tương đối tốt về các lý thuyết cơ bản là điều cần thiết.

- **Chọn mô hình:** điều này sẽ phụ thuộc vào cách trình bày dữ liệu và các ứng dụng. Mô hình quá phức tạp có xu hướng dẫn đến những thách thức trong việc học.
- **Thuật toán học:** có rất nhiều sự thỏa thuận giữa các thuật toán học. Hầu hết các thuật toán sẽ làm việc tốt với các siêu tham số (hyperparameter) đúng để huấn luyện trên một tập hợp dữ liệu cố định cụ thể. Tuy nhiên, việc lựa chọn và điều chỉnh một thuật toán để huấn luyện trên dữ liệu không nhìn thấy yêu cầu một số lượng đáng kể các thử nghiệm.
- **Mạnh mẽ:** Nếu các mô hình, hàm chi phí và thuật toán học được lựa chọn một cách thích hợp, thì ANN sẽ cho kết quả có thể vô cùng mạnh mẽ.

Với việc thực hiện chính xác, ANN có thể được sử dụng một cách tự nhiên học trực tuyến và các ứng dụng tập dữ liệu lớn. Việc thực thi đơn giản của chúng và sự tồn tại chủ yếu là ngoại biên phụ thuộc được thể hiện trong cấu trúc cho phép triển khai nhanh chóng, song song trong phần cứng.

Trong dự báo ANN là cách tiếp cận khá hoàn hảo để dự báo ngắn hạn, do đó tôi sử dụng ANN để dự báo nhu cầu sử dụng trong các lần tiếp theo bằng cách sử dụng những dữ liệu đầu vào được huấn luyện qua từng mốc thời gian thay đổi nhu cầu sử dụng.

Mô hình dự báo nhu cầu sử dụng của ANN sử dụng lý thuyết và các yếu tố ước tính khác để dự báo nhu cầu sử dụng và số lần thay đổi trong tương lai. Họ đã thực hiện truyền ngược với 3 yếu tố ảnh hưởng như biến đầu vào:

- Nhu cầu sử dụng thị trường.
- Nhu cầu sử dụng có trong hệ thống.
- Và nhu cầu sử dụng nhiên liệu.

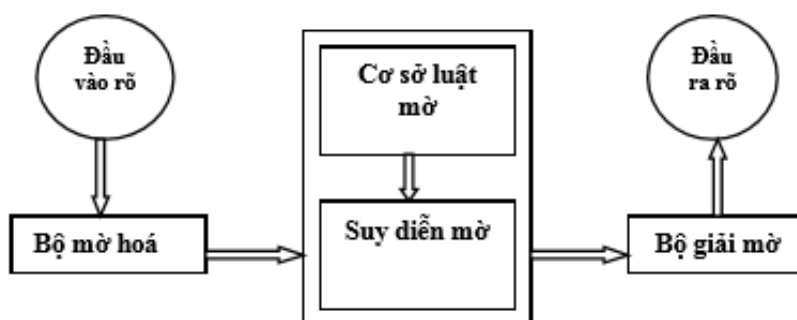
### 2.1.5.2 Máy Vec-tơ hỗ trợ (SVMs)

**Máy vector hỗ trợ (SVM** - viết tắt tên tiếng Anh **support vector machine**) là một khái niệm trong thống kê và khoa học máy tính cho một tập hợp các phương pháp học có giám sát liên quan đến nhau để phân loại và phân tích hồi quy. SVM dạng chuẩn nhận dữ liệu vào và phân loại chúng vào hai lớp khác nhau. Do đó SVM là một thuật toán phân loại nhị phân. Với một bộ các ví dụ luyện tập thuộc hai thể loại cho trước, thuật toán luyện tập SVM xây dựng một mô hình SVM để phân loại các ví dụ khác vào hai thể loại đó. Một mô hình SVM là một cách biểu diễn các điểm trong không gian và lựa chọn ranh giới giữa hai thể loại sao cho khoảng cách từ các ví dụ luyện tập tới ranh giới là xa nhất có thể. Các ví dụ mới cũng được biểu diễn trong cùng một không gian và được thuật toán dự đoán thuộc một trong hai thể loại tùy vào ví dụ đó nằm ở phía nào của ranh giới.

Mô hình SVM để dự báo nhu cầu sử dụng bằng cách sử dụng nhu cầu sử dụng của những lần thay đổi trước đây như một biến đầu vào.

### 2.1.5.3 Suy diễn mờ

Suy diễn hay còn gọi là Logic mờ là một sự khái quát hóa của logic Boolean thường dùng để thiết kế mạch số. Ưu điểm của logic mờ là không cần một mô hình toán học ánh xạ các đầu vào đến đầu ra và không cần các đầu vào chính xác.



**Hình 2.4: Mô hình hệ thống điều khiển mờ**

Phương pháp suy diễn mờ (FIN) để định nhu cầu sử dụng trước. Phương pháp được đề xuất sử dụng FIN để tìm ra các quy tắc mờ với FSOM để đánh nhu cầu sử dụng xác suất của dữ liệu không xác định cho cụm được xác định trước. Kỹ thuật này được sử dụng trong các nhóm mà người sử dụng xác định một cách tự do để đánh giá thuộc tính. FIN có lợi ích không nhất thiết phải có nhóm tương tự để nhóm có thể tự do như mong muốn.



#### 2.1.5.4 Mạng no-ron học sâu

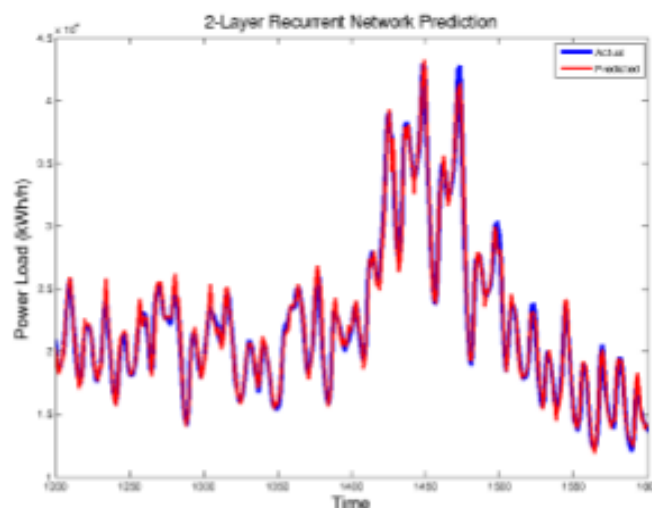
Một số nghiên cứu đã thực hiện thành công các công trình kiến trúc học sâu để dự báo tải điện và thấy rằng kết quả tuyệt vời cho cả hồi quy tuyến tính và hạt nhân cho dữ liệu. Do bộ dữ liệu lớn, họ đã có thể thực hiện các mô hình phi tuyến phức tạp mà không gặp phải nhiều vấn đề quá phù hợp. Nhiều bài báo đã cho rằng những lợi ích của việc huấn luyện không cần phân chia theo lớp để tham gia vào việc khởi tạo mạng lưới sâu.

Khi thử nghiệm với RMSE là 530 kWh/h và tương quan 99,6% với dữ liệu thử nghiệm (xem Bảng 1). Hình 1 cho thấy một so sánh mẫu về nhu cầu tải thực tế có màu xanh (từ bộ dữ liệu thử nghiệm) và dự đoán của mạng hồi quy có màu đỏ. Có nghĩa là tỷ lệ phần trăm sai số tuyệt đối của 0,84-1,56%

Một số kết quả dưới đây cho thấy sự chênh lệch về sai số giữa các phương pháp học.

**Bảng 2.1: Kết quả cho các mô hình học tập khác nhau**

| Phương pháp học | RMSE  | %RMSE |
|-----------------|-------|-------|
| Hồi quy         | 1,540 | 8,3%  |
| NN Tần số       | 1.251 | 6,7%  |
| Hồi quy NN      | 530   | 2,8%  |



**Hình 2.5: Dự báo tải với mạng thần kinh tái diễn**

## **2.2 Dự báo nhu cầu sử dụng điện mặt trời dựa trên học máy**

Kinh tế phát triển nhanh, Việt Nam đang đứng trước sức ép về nhu cầu năng lượng. Tuy đang sử dụng các nguồn tài nguyên truyền thống như than, thủy điện và khí để sản xuất điện, đảm bảo an ninh năng lượng quốc gia, nhưng về lâu dài, Việt Nam đang phải hướng tới xây dựng cơ cấu năng lượng cân bằng thông qua phát triển năng lượng tái tạo, trong đó, năng lượng mặt trời đang có lợi thế khi giá bán thiết bị công nghệ giảm mạnh và xu hướng chuyển sang sử dụng năng lượng tái tạo đang mở rộng trên thế giới. Việt Nam được đánh giá là nước có tiềm năng điện mặt trời rất lớn, tương đương với các nước trong khu vực có thị trường năng lượng mặt trời phát triển như, Trung Quốc, Thái Lan, Philippines hay những thị trường truyền thống như: Ý và Tây Ban Nha. Cụ thể, tổng số giờ nắng của Việt Nam khoảng 1.600 - 2.700 giờ/năm và bức xạ mặt trời bình quân hàng năm đạt khoảng 4 - 5 kWh/m<sup>2</sup>/ngày. Phần nghiên cứu này tập trung vào nghiên cứu hiện tại về dự báo nhu cầu, phân tích các nghiên cứu khác nhau, thiết lập câu hỏi nghiên cứu tổng quát, tính phù hợp và tính khả thi của đề xuất dự án nghiên cứu và khám phá các lĩnh vực cải thiện có thể. Chương này cũng bao gồm nghiên cứu so sánh các kỹ thuật dự báo nhu cầu hiện tại, các mô hình hiện tại. Đồng thời đưa ra phương pháp dự báo nhu cầu sử dụng hồi quy vector hỗ trợ.

### **2.2.1 Mô hình bài toán**

Dự báo nhu cầu sử dụng điện mặt trời đã trở thành một trong những nhiệm vụ quan trọng trong thị trường điện thương mại và quy hoạch. Những hệ quả kinh tế tích cực đã thu hút được nhiều người đầu tư thời gian và tiền bạc để phát triển thuật toán mới để dự đoán nhu cầu sử dụng chính xác. Khía cạnh tài chính này đã thu hút sự chú ý từ nhiều nhà nghiên cứu và đã tạo ra nhiều nghiên cứu và đóng góp quan trọng trong dự báo nhu cầu sử dụng điện mặt trời. Nghiên cứu này góp phần đẩy mạnh hơn nữa trong việc quy hoạch và xây dựng lưới điện thông minh. Khái niệm về kế hoạch nhu cầu sử dụng điện mặt trời linh hoạt một bộ phận của lưới điện thông minh đã mở rộng nghiên cứu về dự báo nhu cầu sử dụng điện mặt trời. Kiến thức chính xác về nhu cầu sử dụng điện mặt trời trong tương lai sẽ giúp người tiêu dùng lên kế hoạch tiêu dùng và nhà cung cấp để lập kế hoạch sản xuất dựa trên hành vi của người sử dụng.

Về mặt kỹ thuật nghiên cứu về dự báo nhu cầu sử dụng điện mặt trời dựa trên nghiên cứu ảnh hưởng của 3 yếu tố sau:

Tổng số giờ chiếu sáng tại các khu vực lắp đặt.

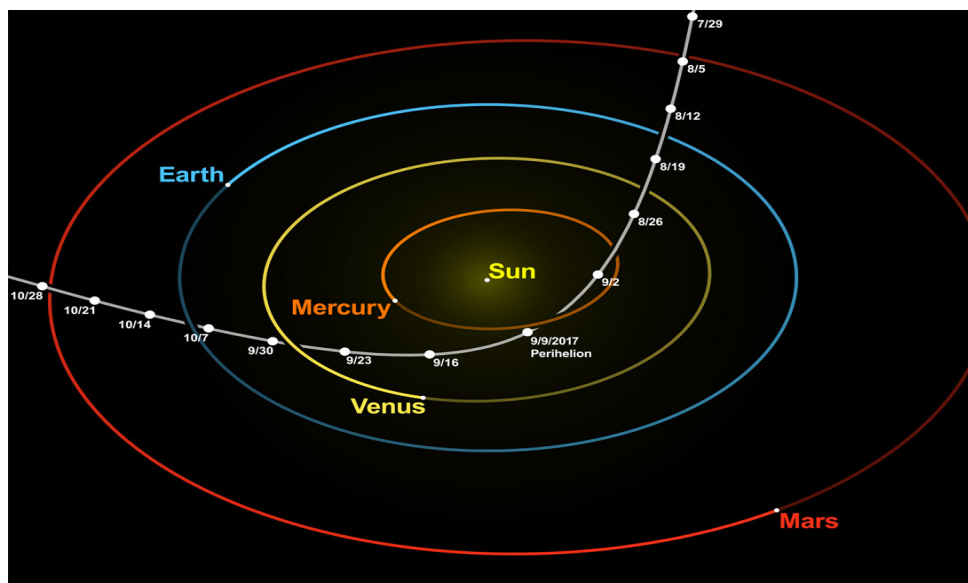
Sản lượng đăng ký lắp đặt với Tập đoàn Điện lực Việt Nam.

Số lượng người lắp đặt hệ thống điện mặt trời theo mốc thời gian.

### **2.2.2 Phân tích các yếu tố ảnh hưởng**

Hiệu suất của tấm PMT, phụ thuộc vào công nghệ chế tạo, về lý thuyết có thể lên tới 30%, tuy nhiên với các sản phẩm thương mại (hiện bán trên thị trường Việt Nam) hiệu suất trung bình đạt được từ 15% đến 18% ở điều kiện tiêu chuẩn của nhà sản xuất. Trong điều kiện thực tế, hiệu suất đạt được chỉ đạt từ 85 - 90% hiệu suất tiêu chuẩn. Bên cạnh đó, một số yếu tố tự nhiên và cách lắp đặt, sử dụng có tác động lớn đến hiệu suất của PMT như: điều kiện khí hậu, hướng lắp đặt, góc nghiêng lắp đặt, chất lượng của các tế bào quang, chế độ bảo dưỡng.

**Cường độ chiếu sáng:** Là thiết bị biến đổi quang năng thành điện năng, vì thế PMT chỉ hoạt động khi có ánh sáng. Khi ánh sáng yếu, mưa, thời tiết âm u, buổi đêm, thì PMT sẽ không hoạt động hoặc với hiệu quả rất thấp. Về góc làm việc của tấm pin NLMT, đối với Việt Nam, hướng Nam là hướng có tổng thời gian đón bức xạ mặt trời nhiều nhất trong năm và là hướng tối ưu để lắp tấm pin năng lượng mặt trời. Để nhận được tối đa năng lượng mặt trời dưới dạng quang năng, khi lắp đặt cần phải quan tâm đến độ nghiêng của thiết bị. Ở Việt Nam góc nghiêng tối ưu nhất khi lắp đặt các tấm pin mặt trời là 15 - 45 độ, độ dốc nghiêng thấp dần về phía Nam. Ví dụ ở Hà Nội, độ nghiêng hợp lý nhất là khoảng 20 - 22 độ, và ở TP HCM là 16 - 18 độ.



Khi nhiệt độ tăng, hiệu suất của tấm pin mặt trời thường giảm một chút.

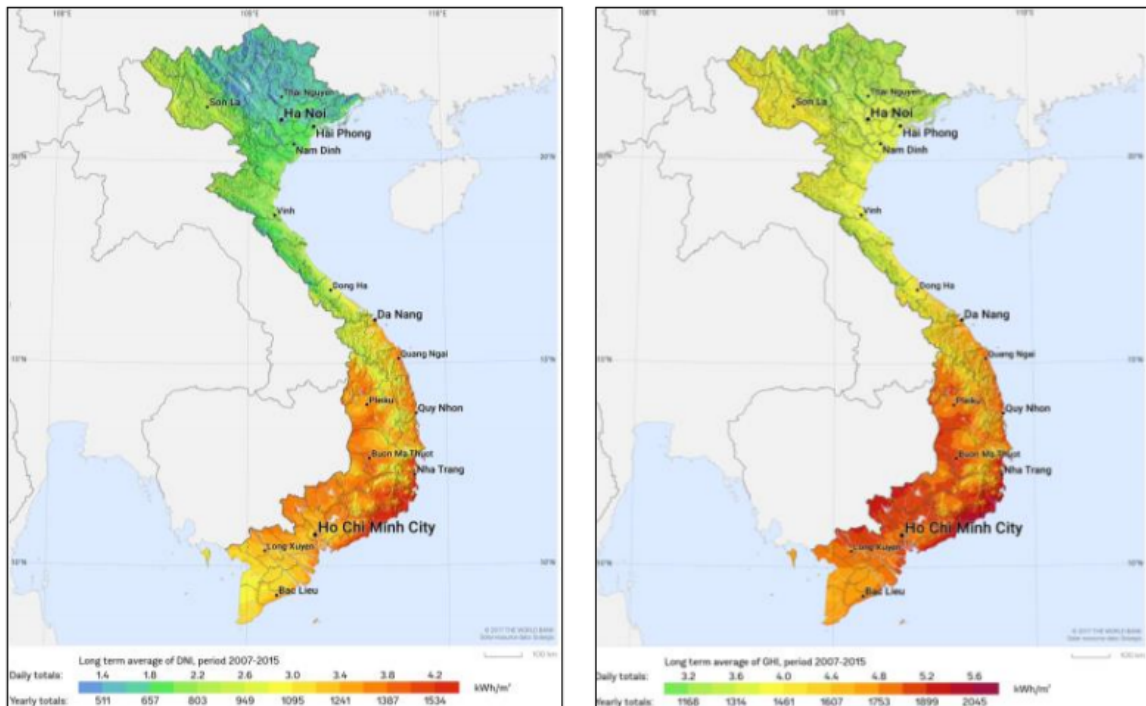


Nhiệt độ cao cũng ảnh hưởng tới pin mặt trời. Các nhà lắp đặt chuyên nghiệp sẽ đảm bảo rằng các tấm pin mặt trời được định vị để nhận đủ lượng luồng gió thổi. Điều này tạo ra sự mát mẻ tự nhiên, giúp giữ tỷ lệ hiệu quả. Khi quy hoạch các dự án điện mặt trời, các yếu tố chính được xem xét khi lựa chọn địa điểm bao gồm:

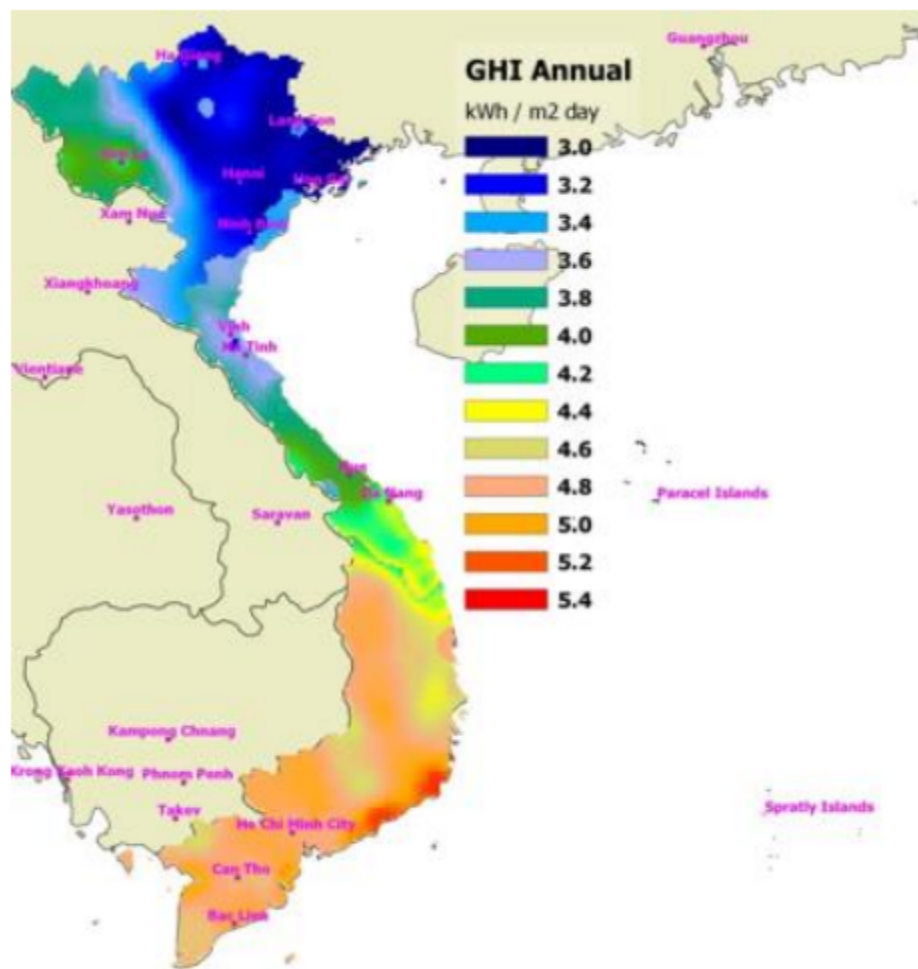
1. Nguồn năng lượng mặt trời
2. Khả năng và chi phí nối lưới
3. Yêu cầu bảo tồn và đa dạng sinh học
4. Quy mô địa điểm, địa hình, lối vào, điều kiện mặt đất
5. Sắp bóng gần và xa
6. Các tác động về cảnh qua, tầm nhìn
7. Hiện trạng sử dụng đất

| TIÊU CHÍ       | LOẠI                           | NGUỒN                                                                                                  | GIẢI ĐOẠN PHÂN TÍCH                  | NHẬN XÉT                                                                                                         |
|----------------|--------------------------------|--------------------------------------------------------------------------------------------------------|--------------------------------------|------------------------------------------------------------------------------------------------------------------|
| Vật lý         | Nguồn bức xạ mặt trời          | SolarGIS                                                                                               | Đánh giá nguồn, tính toán thuộc tính | Sẵn có, tốt nhất và được cập nhật                                                                                |
|                | Cao độ (DEM)                   | Chương trình Tàu con thoi Radar trắc địa                                                               | Đánh giá nguồn, tính toán thuộc tính | Sẵn có tốt nhất                                                                                                  |
|                | Độ dốc                         | SRTM                                                                                                   | Đánh giá nguồn, tính toán thuộc tính | Sẵn có tốt nhất (tính từ DEM)                                                                                    |
| Môi trường     | Sử dụng đất và lớp phủ mặt đất | Dự án lập bản đồ lớp phủ mặt đất thuộc Sáng kiến biến đổi khí hậu của Cơ quan vũ trụ Châu Âu (ESA-CCI) | Đánh giá nguồn, tính toán thuộc tính | Bộ số liệu độ cao toàn cầu; 2015; độ phân giải 300 m; độ chính xác không rõ và chưa được kiểm chứng qua thực địa |
|                | Những khu vực được bảo vệ      | Cơ sở dữ liệu thế giới về những khu vực được bảo vệ                                                    | Đánh giá nguồn                       | Bộ số liệu toàn cầu, không phải tất cả các khu vực đã được các cơ quan của chính phủ thẩm tra                    |
|                | Mặt nước                       | Dự án lập bản đồ lớp phủ mặt đất thuộc Sáng kiến biến đổi khí hậu của Cơ quan vũ trụ Châu Âu           | Đánh giá nguồn, tính toán thuộc tính | Kiểm tra độ chính xác, cần có thông tin thuộc tính tốt hơn (v.d. hồ tự nhiên hay hồ nhân tạo)                    |
| Kinh tế xã hội | Mật độ dân số                  | LandScan                                                                                               | Đánh giá nguồn, tính toán thuộc tính | Bộ dữ liệu toàn cầu, độ phân giải 1 km                                                                           |
|                | Các sân bay, khu vực quân sự.  | GeoViet                                                                                                | Đánh giá nguồn                       |                                                                                                                  |
| Hạ tầng cơ sở  | Đường giao thông, đường điện   | Viện Năng lượng                                                                                        | Tính toán thuộc tính                 | Tốt nhất.                                                                                                        |

Tổng mật độ năng lượng sẽ phụ thuộc và khoảng cách giữa các hàng tấm pin cũng như hiệu suất mô đun của từng tấm pin. Nếu các tấm pin được lắp đặt theo chiều ngang và không có khoảng cách giữa từ mô đun, tổng mật độ năng lượng sẽ bằng mật độ năng lượng mô đun (100–150 MWp/km<sup>2</sup> đối với mô đun silicon). Xem xét một số dự án quy mô lớn và thảo luận với một số đơn vị lắp đặt cho thấy khoảng cách tối thiểu đối là khoảng 3,5m giữa các hàng và có thể là 4 – 5m trên thực tế. Thực tiễn hiện nay, theo quy định của Bộ Công Thương yêu cầu 1,2 – 1,3 ha/MWp, tương đương với 77-82 MWp/km<sup>2</sup>. Giả định rằng tỉ lệ DC/AC = 1,35, chúng ta có thể xác định mật độ năng lượng mặt trời khoảng 55 – 60MW/km<sup>2</sup> (AC) đối với hệ thống pin cố định. Xem xét nhiều thông số khác nhau về độ dốc, bề mặt đất.... Từ vận sử dụng thông số mật độ năng lượng trung bình là 50 MW/km<sup>2</sup>.  
 Thống kê công suất theo khu vực dựa trên bản đồ cường độ chiếu sáng



**Hình 2.6. Bản đồ cường độ chiếu sáng**



**Hình 2.7. Tiềm năng điện mặt trời tại Việt Nam**

Sản lượng đăng ký với Tập đoàn điện lực tính đến tháng 8/2018

| STT        | Tên khách hàng                     | Số khách hàng | Công suất dự án (kWp) | Ngày lắp đặt công tơ 2 chiều | Sản lượng phát lên lưới (tỷ lệ đến 31/7/2018) (kWh) |
|------------|------------------------------------|---------------|-----------------------|------------------------------|-----------------------------------------------------|
|            | <b>5 TCTĐL</b>                     | <b>540</b>    | <b>9.717,98</b>       |                              |                                                     |
| <b>I</b>   | <b>Tổng Công ty ĐL miền Bắc</b>    | <b>1</b>      | <b>90,00</b>          |                              |                                                     |
|            | Công ty Điện lực Bắc Giang         | 1             |                       |                              |                                                     |
| 1          | Công ty TNHH Vina Solar Technology |               | 90,00                 |                              |                                                     |
| <b>II</b>  | <b>Tổng Công ty ĐL miền Nam</b>    | <b>168,00</b> | <b>4.420,46</b>       | <b>22/08/2018</b>            |                                                     |
| <b>III</b> | <b>Tổng Công ty ĐL miền Trung</b>  | <b>104</b>    | <b>1766,89</b>        |                              | <b>154.835,64</b>                                   |
|            | <b>Công ty Điện lực Quảng Bình</b> | <b>3</b>      | <b>14,00</b>          |                              | <b>466,4</b>                                        |
| 1          | Thái Hồng Quân                     |               | 5,00                  | 28/06/2016                   | 113,0                                               |
| 2          | Lê Đức Dũng                        |               | 2,00                  | 07/02/2018                   | 95,0                                                |
| 3          | Dương Quang Thành                  |               | 5,00                  | 29/10/2017                   | 253,0                                               |
| 4          | Nguyễn Phú Thám                    |               | 2,00                  | 21/07/2018                   | 5,4                                                 |
|            | <b>Công ty Điện lực Đà Nẵng</b>    | <b>33</b>     | <b>171,03</b>         |                              | <b>20.489,8</b>                                     |
| 1          | Lê Ngọc Nam                        |               | 2,00                  | 17/03/2016                   | 194,0                                               |
| 2          | Nguyễn Thành                       |               | 5,00                  | 18/10/2017                   | 2.482,0                                             |
| 3          | Trần Quốc Huy                      |               | 3,00                  | 20/03/2018                   | 446,0                                               |
| 4          | Phan Thị Hiền                      |               | 1,00                  | 24/08/2016                   | 671,0                                               |
| 5          | Nguyễn Thị Kim Hồng                |               | 6,00                  | 04/12/2017                   | 1.199,0                                             |
| 6          | Trần Đình Nhân                     |               | 5,00                  | 04/07/2017                   | 3.443,0                                             |
| 7          | Lưu Bình Thanh                     |               | 2,00                  | 17/04/2018                   | 51,0                                                |
| 8          | Ngô Tấn Cư                         |               | 2,00                  | 30/05/2018                   | 55,0                                                |
| 9          | Trần Thế Lan                       |               | 2,00                  | 23/05/2018                   | 56,0                                                |
| 10         | Công ty TNHH kỹ thuật PO SO        |               | 1,96                  | 07/05/2018                   | 3,0                                                 |
| 11         | Nguyễn Gia Phước                   |               | 2,85                  | 23/05/2018                   | 371,0                                               |
| 12         | Nguyễn Thanh An                    |               | 2,64                  | 26/07/2018                   | -                                                   |
| 13         | Ngô Anh Đồng                       |               | 1,00                  | 27/06/2018                   | 71,0                                                |

| STT | Tên khách hàng                                                        | Số khách hàng | Công suất dự án (kWp) | Ngày lắp đặt công tơ 2 chiều | Sản lượng phát lên lưới (tỷ lệ đến 31/7/2018) (kWh) |
|-----|-----------------------------------------------------------------------|---------------|-----------------------|------------------------------|-----------------------------------------------------|
| 14  | Công ty Lưới điện Cao Thế Miền Trung-Tổng công ty Điện lực Miền Trung |               | 35,00                 | 13/02/2018                   | 1.297,0                                             |
| 15  | Võ Như Vũ                                                             |               | 1,32                  | 19/05/2018                   | 4,0                                                 |
| 16  | Lê Văn Đoàn                                                           |               | 2,50                  | 19/05/2018                   | 90,0                                                |
| 17  | Huỳnh Tấn Thành                                                       |               | 3,00                  | 27/06/2016                   | 620,0                                               |
| 18  | Trần Đình Lợi                                                         |               | 3,00                  | 17/10/2017                   | 874,0                                               |
| 19  | Nguyễn Minh Vũ                                                        |               | 3,00                  | 10/03/2018                   | 1.597,0                                             |
| 20  | Nguyễn Văn Thông                                                      |               | 4,00                  | 26/03/2018                   | 258,0                                               |
| 21  | Cty CP Nông nghiệp CNC AFARM                                          |               | 15,00                 | 23/03/2018                   | 995,0                                               |
| 22  | Huỳnh Phước Hoàng                                                     |               | 3,00                  | 25/05/2018                   | 159,0                                               |
| 23  | Vũ Giai Tuyên                                                         |               | 3,00                  | 01/06/2016                   | 422,0                                               |
| 24  | Ung Thế Anh                                                           |               | 3,00                  | 03/07/2018                   | 294,0                                               |
| 25  | Nguyễn Văn Tuấn                                                       |               | 2,74                  | 05/10/2016                   | 150,0                                               |
| 26  | Đặng Thị Mỹ Lợi                                                       |               | 2,40                  | 21/8/2017                    | 992,0                                               |
| 27  | Hồ Thị Ngân                                                           |               | 3,00                  | 14/10/2017                   | 130,0                                               |
| 28  | Ngô Hoàng Đại Việt                                                    |               | 2,90                  | 02/4/2018                    | 69,8                                                |
| 29  | Trần Thị Thám                                                         |               | 2,04                  | 13/4/2018                    | 75,0                                                |
| 30  | Hộ KD Bida Thông Nhất                                                 |               | 22,00                 | 21/04/2018                   | 3.200,0                                             |
| 31  | Ngô Việt Linh                                                         |               | 1,08                  | 12/05/2018                   | 35,0                                                |
| 32  | Võ Đức Hiếu                                                           |               | 1,60                  | 01/06/2018                   | 186,0                                               |
| 33  | Cty TNHH MTV ĐLĐN                                                     |               | 21,00                 | 2017                         | -                                                   |
|     | <b>Công ty Điện lực Quảng Nam</b>                                     | <b>3</b>      | <b>140,00</b>         |                              | <b>5.124,0</b>                                      |
| 1   | Trung tâm văn hóa thể thao Hội An                                     |               | 50,00                 | 12/01/2018                   | 1.840,0                                             |
| 2   | Trường THPT Lê Quý Đôn                                                |               | 10,00                 | 22/04/2018                   | 3.181,0                                             |
| 3   | Chi nhánh Tập đoàn dệt may Việt Nam                                   |               | 80,00                 | 23/05/2017                   | 103,0                                               |
|     | <b>Công ty Điện lực Quảng Ngãi</b>                                    | <b>3</b>      | <b>18,00</b>          |                              | <b>2.267,0</b>                                      |
| 1   | Trung tâm nuôi dạy trẻ khuyết tật Võ Hồng Sơn                         |               | 10,00                 | 27/10/2017                   | 175,0                                               |
| 2   | Trần Đức Thịnh (Nhà thờ)                                              |               | 3,00                  | 20/10/2017                   | 1.984,0                                             |
| 3   | Cao Minh Tuấn                                                         |               | 5,00                  | 02/02/2018                   | 108,0                                               |



| STT | Tên khách hàng                                | Số khách hàng | Công suất dự án (kWp) | Ngày lắp đặt công tơ 2 chiều | Sản lượng phát lên lưới (tỷ lệ đến 31/7/2018) (kWh) |
|-----|-----------------------------------------------|---------------|-----------------------|------------------------------|-----------------------------------------------------|
|     | <b>Công ty Điện lực Bình Định</b>             | <b>10</b>     | <b>40,66</b>          |                              | <b>2.254,0</b>                                      |
| 1   | Đặng Công Thuận                               |               | 1,00                  | 14/12/2017                   | 285,0                                               |
| 2   | Nguyễn Thị Nga                                |               | 5,00                  | 02/10/2017                   | 4,0                                                 |
| 3   | DNTT KHÁCH SẠN KHANG KHANG                    |               | 5,00                  | 23/02/2017                   | 5,0                                                 |
| 4   | Nguyễn Thanh Giang                            |               | 5,20                  | 20/04/2018                   | -                                                   |
| 5   | Phan Thị Mĩ Mĩ                                |               | 3,60                  | 27/06/2018                   | 100,0                                               |
| 6   | Trần Đình Tấn                                 |               | 3,20                  | 05/07/2018                   | 755,0                                               |
| 7   | Nguyễn Thị Bảy                                |               | 3,96                  | 22/05/2018                   | 698,0                                               |
| 8   | Lê Quang Hải                                  |               | 3,30                  | 14/06/2018                   | 285,0                                               |
| 9   | Đặng Minh Công                                |               | 2,40                  | 20/06/2018                   | 122,0                                               |
| 10  | Công ty Cổ phần thủy điện Trà Xom             |               | 10,00                 |                              |                                                     |
|     | <b>Công ty Điện lực Phú Yên</b>               | <b>16</b>     | <b>493,00</b>         |                              | <b>25.669,0</b>                                     |
| 1   | Công ty TNHH Xây dựng & Quảng cáo Phương Tuấn |               | 60,00                 |                              |                                                     |
| 2   | Nguyễn Thị Nhơn                               |               | 10,00                 |                              |                                                     |
| 3   | Công ty Điện lực Phú Yên (Văn phòng Công ty)  |               | 40,00                 | 08/02/2018                   | 25.322,0                                            |
| 4   | Công ty Điện lực Phú Yên (ĐL Tuy Hòa)         |               | 35,00                 |                              |                                                     |
| 5   | Công ty Điện lực Phú Yên (ĐL Đông Hòa)        |               | 30,00                 |                              |                                                     |
| 6   | Công ty Điện lực Phú Yên (ĐL Phú Hòa)         |               | 30,00                 |                              |                                                     |
| 7   | Công ty Điện lực Phú Yên (ĐL Tây Hòa)         |               | 50,00                 |                              |                                                     |
| 8   | Công ty Điện lực Phú Yên (ĐL Sơn Hòa)         |               | 40,00                 |                              |                                                     |
| 9   | Công ty Điện lực Phú Yên (ĐL Sông Hinh)       |               | 50,00                 |                              |                                                     |
| 10  | Công ty Điện lực Phú Yên (ĐL Tuy An)          |               | 40,00                 |                              |                                                     |
| 11  | Công ty Điện lực Phú Yên (ĐL Đông Xuân)       |               | 50,00                 |                              |                                                     |
| 12  | Công ty Điện lực Phú Yên (ĐL Sông Cầu)        |               | 35,00                 |                              |                                                     |
| 13  | DNTN Xăng dầu Kiều Mỹ                         |               | 2,00                  |                              |                                                     |
| 14  | Nguyễn Anh Tuấn                               |               | 2,00                  |                              |                                                     |
| 15  | DNTN Hà Mai                                   |               | 16,00                 | 20/7/2018                    | 347,0                                               |
| 16  | Đoàn Đình Dương                               |               | 3,00                  |                              |                                                     |

| STT | Tên khách hàng                       | Số khách hàng | Công suất dự án (kWp) | Ngày lắp đặt công tơ 2 chiều | Sản lượng phát lên lưới (tỷ lệ đến 31/7/2018) (kWh) |
|-----|--------------------------------------|---------------|-----------------------|------------------------------|-----------------------------------------------------|
|     | <b>Công ty Điện lực Khánh Hòa</b>    | <b>16</b>     | <b>132,70</b>         |                              | <b>4.906,0</b>                                      |
| 1   | Nguyễn Thanh Lâm                     |               | 10,00                 | 24/01/2018                   | 1.209,0                                             |
| 2   | Nguyễn Thị Khai                      |               | 10,00                 | 15/01/2018                   | -                                                   |
| 3   | Nguyễn Hưng                          |               | 10,00                 | 24/01/2018                   | -                                                   |
| 4   | Nguyễn Thị Lưu                       |               | 10,00                 | 18/04/2018                   | 2.100,0                                             |
| 5   | Trần Bảo Phương                      |               | 12,00                 | 15/6/2018                    | -                                                   |
| 6   | Nguyễn Thị Nhung                     |               | 15,00                 | 30/07/048                    | -                                                   |
| 7   | Nguyễn Cao Kỳ                        |               | 10,20                 | 43284                        | 11,0                                                |
| 8   | Công Ty TNHH Viễn Thông Hoàng Ngọc   |               | 10,20                 | 43286                        | -                                                   |
| 9   | Lê Thị Hạnh                          |               | 5,10                  | 43286                        | -                                                   |
| 10  | Hộ Kinh Doanh Pizza Giang            |               | 10,20                 | 18/01/2018                   | -                                                   |
| 11  | Nguyễn Chiến Thắng                   |               | 10,00                 | 18/01/2018                   | 1.109,0                                             |
| 12  | Nguyễn Hữu Khánh                     |               | 3,00                  | 43264                        | 211,0                                               |
| 13  | Trương Kinh                          |               | 3,00                  | 43264                        | 266,0                                               |
| 14  | Lê Xuân Cường                        |               | 10,00                 | 43290                        | -                                                   |
| 15  | Nguyễn Thị Mười                      |               | 2,00                  | 43317                        | -                                                   |
| 16  | Nguyễn Thanh Huy                     |               | 2,00                  | 28/8/2017                    | -                                                   |
|     | <b>Công ty Điện lực Gia Lai</b>      | <b>9</b>      | <b>94,5</b>           |                              | <b>314,00</b>                                       |
| 1   | Công ty Cổ phần Điện Gia Lai         |               | 20,00                 | 26/01/2018                   | 85,00                                               |
| 2   | Lê Tuấn Thành                        |               | 6,40                  |                              |                                                     |
| 3   | Nguyễn Đình Việt                     |               | 0,60                  |                              |                                                     |
| 4   | Nguyễn Văn Quang                     |               | 5,50                  |                              |                                                     |
| 5   | Lê Thị Thủy                          |               | 0,60                  |                              |                                                     |
| 6   | Công ty TNHH Tư Vấn thiết kế Thiên Ý |               | 3,40                  |                              |                                                     |
| 7   | Lê Thị Tâm                           |               | 5,00                  | 10/06/2018                   | 6,00                                                |
| 8   | Nguyễn Thanh Thảo                    |               | 3,00                  |                              |                                                     |
| 9   | Công ty ĐL Gia Lai                   |               | 50,00                 | 15/04/2018                   | 223,00                                              |
|     | <b>Công ty Điện lực Kon Tum</b>      | <b>1</b>      | <b>35,00</b>          |                              | <b>4,44</b>                                         |
| 1   | Công ty ĐL Kon Tum                   |               | 35,00                 | 06/02/2018                   | 4,44                                                |



| STT       | Tên khách hàng                        | Số khách hàng | Công suất dự<br>án (kWp) | Ngày lắp đặt công<br>tơ 2 chiều | Sản lượng<br>phát lên lưới<br>(tính đến<br>31/7/2018)<br>(kWh) |
|-----------|---------------------------------------|---------------|--------------------------|---------------------------------|----------------------------------------------------------------|
|           | <b>Công ty Điện lực Đắk Lắk</b>       | <b>8</b>      | <b>588,00</b>            |                                 | <b>85960,00</b>                                                |
| 1         | Y Giang Niê Knong                     |               | 2,00                     | 01/2017                         |                                                                |
| 2         | Phạm Thị Tuất                         |               | 1,00                     | 09/2017                         |                                                                |
| 3         | Nguyễn Văn Thân                       |               | 2,00                     | 04/2017                         |                                                                |
| 4         | Nguyễn Thị Quý                        |               | 1,00                     | 08/2017                         |                                                                |
| 5         | Y Phu E Ban                           |               | 2,00                     | 04/2017                         |                                                                |
| 6         | Cty đầu tư XD cho CBMT                |               | 438,00                   | 01/2018                         | 58700,00                                                       |
| 7         | Cty Điện lực Đắk Lắk                  |               | 56,00                    | 02/2018                         | 184,00                                                         |
| 8         | Cty CP sách - thiết bị trường học     |               | 86,00                    | 01/2018                         | 27076,00                                                       |
|           | <b>Công ty Điện lực Đắk Nông</b>      | <b>2</b>      | <b>40,00</b>             |                                 | <b>7381,00</b>                                                 |
| 1         | Công ty Điện lực Đắk Nông             |               | 35,00                    | 25/2/2018                       | 3824,00                                                        |
| 2         | Trung tâm quỹ đất huyện Cujut         |               | 5,00                     | 01/09/2017                      | 3557,00                                                        |
| <b>II</b> | <b>Tổng Công ty ĐL TP Hồ Chí Minh</b> | <b>267</b>    | <b>3440,63</b>           |                                 |                                                                |
|           | <b>Công ty Điện lực Sài Gòn</b>       | <b>15</b>     | <b>121,20</b>            |                                 |                                                                |
| 1         | Nguyễn Đức Thọ                        |               | 3,00                     | 31/01/2018                      |                                                                |
| 2         | Võ Việt Lập                           |               | 3,00                     | 26/12/2017                      |                                                                |
| 3         | Trụ sở Tổng công ty Điện lực TP.HCM   |               | 28,60                    | 17/6/2017                       |                                                                |
| 4         | Đặng Thị Kim Vinh                     |               | 5,00                     | 04/02/2018                      |                                                                |
| 5         | Dương Vĩnh Phúc                       |               | 3,00                     | 26/12/2017                      |                                                                |
| 6         | Công ty CP VLXD Sư Tử Biển            |               | 5,00                     | 26/12/2017                      |                                                                |
| 7         | NGUYỄN THUY CHUNG                     |               | 2,50                     | 26/12/2017                      |                                                                |
| 8         | Thái Đôn Thành                        |               | 9,60                     | 06/3/2018                       |                                                                |
| 9         | Lưu Thế Khanh                         |               | 3,00                     | 14/3/2018                       |                                                                |
| 10        | Văn phòng Tổng Giám mục Việt Nam      |               | 30,00                    |                                 |                                                                |
| 11        | Nguyễn Thị Kim Cúc - Nguyễn Tường Vũ  |               | 3,00                     |                                 |                                                                |
| 12        | Nguyễn Thị Kim Dung                   |               | 15,00                    |                                 |                                                                |
| 13        | Nguyễn Văn A                          |               | 5,00                     |                                 |                                                                |
| 14        | Nguyễn Thục Trâm                      |               | 3,00                     |                                 |                                                                |
| 15        | LÂM THẾ KHAN                          |               | 2,50                     |                                 |                                                                |

| STT | Tên khách hàng                      | Số khách hàng | Công suất dự<br>án (kWp) | Ngày lắp đặt công<br>tơ 2 chiều | Sản lượng<br>phát lên lưới<br>(tính đến<br>31/7/2018)<br>(kWh) |
|-----|-------------------------------------|---------------|--------------------------|---------------------------------|----------------------------------------------------------------|
|     | <b>Công ty Điện lực Phú Thọ</b>     | <b>11</b>     | <b>185,00</b>            |                                 |                                                                |
| 1   | Trường Đại học Bách Khoa            |               | 62,00                    | 30/9/2017                       |                                                                |
| 2   | Trường Đại học Bách Khoa            |               | 88,00                    | 04/10/2017                      |                                                                |
| 3   | Nguyễn Văn An                       |               | 3,00                     | 10/11/2017                      |                                                                |
| 4   | Phan Thị Thu Hương                  |               | 4,00                     | 15/11/2017                      |                                                                |
| 5   | Hoàng Trung Nam                     |               | 4,00                     | 07/12/2017                      |                                                                |
| 6   | Lê Minh Nam                         |               | 4,00                     | 07/12/2017                      |                                                                |
| 7   | Hoàng Thảo                          |               | 4,00                     | 15/11/2017                      |                                                                |
| 8   | Lâm Văn Giang                       |               | 4,00                     | 25/12/2017                      |                                                                |
| 9   | Đào Duy Khanh                       |               | 4,00                     | 09/01/2018                      |                                                                |
| 10  | Nguyễn Hoàng Duy                    |               | 4,00                     | 13/02/2018                      |                                                                |
| 11  | Hộ kinh doanh Gaming                |               | 4,00                     | 28/03/2018                      |                                                                |
|     | <b>Công ty Điện lực An Phú Đông</b> | <b>19</b>     | <b>1.154,40</b>          |                                 |                                                                |
| 1   | Vương Trí Dũng                      |               | 10,00                    | 29/12/2017                      |                                                                |
| 2   | Cty TNHH TM Lê Chánh Nguyên         |               | 3,60                     | 27/9/2017                       |                                                                |
| 3   | Ủy ban nhân dân Quận 12             |               | 20,00                    | 24/6/2017                       |                                                                |
| 4   | Công ty TNHH Kỹ Thuật Đạt           |               | 12,00                    | 11/10/2017                      |                                                                |
| 5   | Nguyễn Văn Ước                      |               | 5,00                     | 29/8/2017                       |                                                                |
| 6   | Cty TNHH CNĐT-VT Quốc tế Đồng Dương |               | 5,00                     | 12/10/2017                      |                                                                |
| 7   | Cty CP ĐT XD và TM Phú Điện         |               | 980,00                   | 19/10/2017                      |                                                                |
| 8   | Cty TNHH SX TM DV Tín Phát          |               | 30,00                    | 24/11/2017                      |                                                                |
| 9   | Trần Văn Dũng                       |               | 3,00                     | 23/11/2017                      |                                                                |
| 10  | Cty Điện lực An Phú Đông            |               | 48,00                    | 09/11/2017                      |                                                                |
| 11  | Lê Văn Ngọc                         |               | 2,40                     | 01/12/2017                      |                                                                |
| 12  | Lê Tiên Thiên                       |               | 5,40                     | 24/11/2017                      |                                                                |
| 13  | Cty TNHH Kỹ thuật Năng Lượng Xanh   |               | 5,00                     | 13/02/2018                      |                                                                |
| 14  | Cty Cổ phần SOUTH SOLAR             |               | 6,00                     | 27/02/2018                      |                                                                |
| 15  | Hồng Phương                         |               | 1,00                     | 13/02/2018                      |                                                                |
| 16  | Hồ Nguyễn Thủy Vy                   |               | 3,00                     | 08/3/2018                       |                                                                |

| STT | Tên khách hàng                                                        | Số khách hàng | Công suất dự án (kWp) | Ngày lắp đặt công tơ 2 chiều | Sản lượng phát lên lưới (tính kể đến 31/7/2018) (kWh) |
|-----|-----------------------------------------------------------------------|---------------|-----------------------|------------------------------|-------------------------------------------------------|
| 16  | Hồ Nguyễn Thủy Vy                                                     |               | 3,00                  | 08/3/2018                    |                                                       |
| 17  | BCH Quân sự Quận 12                                                   |               | 7,00                  |                              |                                                       |
| 18  | Phạm Văn Thối                                                         |               | 1,00                  |                              |                                                       |
| 19  | Nguyễn Văn Căn                                                        |               | 7,00                  |                              |                                                       |
|     | <b>Công ty Điện lực Chợ Lớn</b>                                       | <b>10</b>     | <b>47,80</b>          |                              |                                                       |
| 1   | Trình Phương Thành                                                    |               | 3,00                  | 19/12/2017                   |                                                       |
| 2   | Phạm Thị Thu Dung                                                     |               | 3,00                  | 13/10/2017                   |                                                       |
| 3   | Đặng Chuẩn                                                            |               | 3,30                  | 21/11/2017                   |                                                       |
| 4   | Nguyễn Quang Tính                                                     |               | 3,00                  | 21/11/2017                   |                                                       |
| 5   | Nguyễn Xuân Hải                                                       |               | 5,00                  | 01/12/2017                   |                                                       |
| 6   | Nguyễn Xuân Hải                                                       |               | 10,00                 | 01/12/2017                   |                                                       |
| 7   | Công ty TNHH TM VÀ KT Tân Kỳ Nguyên                                   |               | 10,00                 | 27/3/2018                    |                                                       |
| 8   | Đinh Thanh Hương                                                      |               | 3,00                  | 28/3/2018                    |                                                       |
| 9   | Đinh Xuân Giao                                                        |               | 4,50                  | 23/3/2018                    |                                                       |
| 10  | Thăm mý viện Bắc Sỹ Minh                                              |               | 3,00                  | 19/3/2018                    |                                                       |
|     | <b>Công ty Điện lực Gia Định</b>                                      | <b>23</b>     | <b>253,66</b>         |                              |                                                       |
| 1   | Nguyễn Xuân Chi                                                       |               | 15,00                 | 27/10/2017                   |                                                       |
| 2   | Nguyễn Minh Tú                                                        |               | 3,00                  | 06/3/2018                    |                                                       |
| 3   | Hứa Văn Thế Phúc                                                      |               | 1,20                  | 21/12/2017                   |                                                       |
| 4   | Công ty TNHH Hoàng Hà                                                 |               | 15,00                 | 22/6/2017                    |                                                       |
| 5   | Công ty CP in Thanh Hiền                                              |               | 21,00                 | 19/12/2017                   |                                                       |
| 6   | Phạm Thị Thanh Vân                                                    |               | 3,00                  |                              |                                                       |
| 7   | Trụ sở Phòng Thông tin tuyên truyền An toàn - Tiết kiệm năng lượng cũ |               | 25,00                 | 19/11/2016                   |                                                       |
| 8   | Đại học Văn Lang                                                      |               | 55,00                 | 11/7/2017                    |                                                       |
| 9   | Nguyễn Quốc Tiên                                                      |               | 6,00                  | 31/10/2017                   |                                                       |
| 10  | Lê Công Huân                                                          |               | 2,00                  |                              |                                                       |
| 11  | Lê Quang Huy                                                          |               | 1,00                  |                              |                                                       |
| 12  | Cty CPĐT XD Cát Tường                                                 |               | 13,00                 | 07/12/2017                   |                                                       |
| 13  | Hồ Việt Hải                                                           |               | 5,00                  | 31/10/2017                   |                                                       |

## 2.3 Ứng dụng máy Vec-tơ hỗ trợ (SVMs) trong bài toán dự báo nhu cầu sử dụng điện mặt trời

Trong những năm gần đây, máy vector hỗ trợ (SVM) đã được sử dụng thành công trên phân loại nhiệm vụ trong các lĩnh vực ứng dụng rất khác nhau. Vì trọng tâm của công việc này nằm trên dự báo năng lượng PV, và do đó các giá trị từ không gian liên tục, một phần mở rộng của SVM gọi là hồi quy vector hỗ trợ (SVR) được áp dụng. Khái niệm chính của các phương pháp này vẫn giữ nguyên: Tìm một hàm đại diện cho một tập hợp dữ liệu đào tạo nhất định và áp dụng chức năng này cho các điểm dữ liệu chưa từng thấy trước đó để có được dự đoán dựa trên dữ liệu đào tạo.

### 2.3.1 Máy Vec-tơ hỗ trợ (SVMs)

Trong tất cả các kịch bản học tập, có ba bộ dữ liệu quan trọng: tập huấn luyện, bài kiểm tra thiết lập và bộ đánh giá. Trong khi tập huấn luyện được sử dụng để huấn luyện thuật toán học tập, tham số và tối ưu hóa mô hình được thực hiện trên tập kiểm tra. Sau đó, tối ưu hóa mô hình được phân tích trên một tập dữ liệu đánh

giá độc lập. Trong trường hợp này, tập huấn luyện bao gồm các phép đo và dự báo cho một số ngày được xác định trước trước timestep thực tế được dự báo.

Mục đích của thuật toán này là tìm hàm dự đoán  $f: X \rightarrow Y$  gán a nhãn y cho mẫu trước khi tìm được x. Do đó, hàm f được học trên một tập huấn luyện

$$\mathcal{T} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\} \subset \mathcal{X} \times \mathcal{Y}, \quad (2.1)$$

Trong đó N biểu thị kích thước huấn luyện có chứa các mối quan hệ đã biết giữa các mẫu và nhãn. Sau khi giới thiệu bộ huấn luyện, bây giờ có thể phân biệt giữa học có giám sát và không giám sát.

**Học có giám sát:** Trong trường hợp học có giám sát, sự liên kết các mẫu và nhãn trong T có sẵn cho thuật toán học tập như đã thấy trong phương trình (2.1). Nổi bật nhất của nhóm học có giám sát là phân loại và hồi quy. Dành cho nhiệm vụ phân loại, đầu ra được đưa ra là một tập hữu hạn các giá trị rời rạc gọi là các lớp, ví dụ: đối với giá trị nhị phân được sử dụng trong các phần như sau  $Y = \{1, -1\}$ . Ngược lại, trong hầu hết các trường hợp mà một hàm hồi quy được học, không gian đầu ra bao gồm các nhãn có giá trị thực liên tục. Trong việc này, chủ yếu là các hàm hồi quy được học, do đó

$\mathbf{x} \in X = \mathbb{R}$  cho các mẫu và  $y \in Y = \mathbb{R}$  cho nhãn áp dụng.

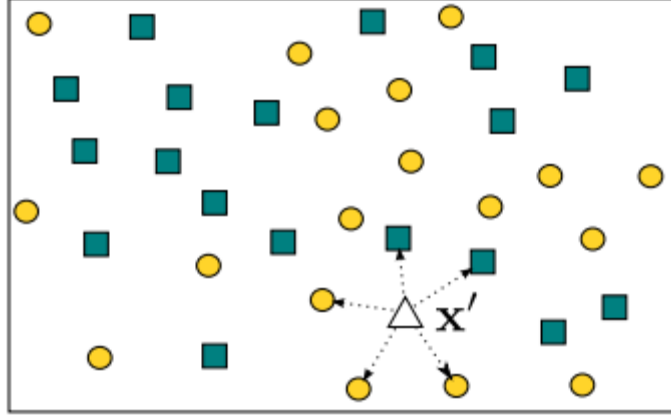
Để tạo thuận lợi cho việc hiểu và sử dụng các mô hình phân loại và hồi quy, k-láng giềng ( $k$ -NN), thường được gọi là ví dụ đơn giản cho người được giám sát thuật toán học máy. Cách tiếp cận này sử dụng các nhãn của các mẫu lân cận quyết định nhãn nào sẽ được gán cho mẫu mới x. Thuật toán xem xét k điểm dữ liệu gần nhất trong không gian đầu vào và sử dụng nhãn mà phần lớn các mẫu giữ. Hàm quyết định phân loại  $f_{knn}$  với không gian đầu ra  $Y = \{1, -1\}$  có thể được định nghĩa là

$$f_{kNN}(\mathbf{x}') = \begin{cases} 1 & \text{if } \sum_{i \in \mathcal{N}_k(\mathbf{x}')} y_i > 0 \\ -1 & \text{if } \sum_{i \in \mathcal{N}_k(\mathbf{x}')} y_i \leq 0, \end{cases} \quad (2.2)$$

Trong đó  $N_k$  biểu thị tập hợp các láng giềng gần nhất của một mẫu mới x' trong đào tạo T. Để đo khoảng cách không gian đầu vào khoảng cách Euclidian.

$$\delta(\mathbf{x}, \mathbf{x}') = \left( \sum_{j=1}^d (x_j - x'_j)^2 \right)^{\frac{1}{2}} \quad (2.3)$$

Thường được áp dụng trong hình 2.1 là một ví dụ minh họa quá trình ra quyết định của  $k$ -NN khi một mẫu mới được đưa vào một tập hợp các điểm dữ liệu đã được dán nhãn.



**Hình 2.8: Ví dụ minh họa về phân loại  $k$ - láng giềng gần nhất.**

Hai lớp màu vàng vòng tròn và hình vuông xanh lá cây) được cung cấp một điểm dữ liệu mới  $x'$  được giới thiệu. Đối với  $k=5$  là nhãn của  $x'$  sẽ là một vòng tròn vàng khi ba trong năm láng giềng gần nhất được dán nhãn theo cách này.

Để chứng minh kết nối giữa phân loại và hồi quy, mô hình  $k$ -NN cho hồi quy được giới thiệu như sau. Thuật toán  $k$ -NN dễ dàng được mở rộng để sử dụng một hàm hồi quy bằng cách thay thế  $f_{knn}$  từ phương trình trong hình 2.2 bằng trung bình số học và tính trung bình của các nhãn láng giềng gần nhất. Trong việc này khoảng cách là phần có trọng số của Bailey và Jain.

$$f_{wkNN}(\mathbf{x}') = \sum_{i \in \mathcal{N}_K(\mathbf{x}')} \frac{\delta(\mathbf{x}_i, \mathbf{x}')^{-1}}{\sum_{j \in \mathcal{N}_K(\mathbf{x}')} \delta(\mathbf{x}_j, \mathbf{x}')^{-1}} y_i, \quad (2.4)$$

Nơi nghịch đảo của khoảng cách Euclidian, được sử dụng để xác định phần quan trọng của nhãn láng giềng tùy thuộc vào khoảng cách của nó đến mẫu mới  $x'$

Với cụm phân loại các mẫu không có nhãn từ trước  $y$  có thể là các mẫu có cùng điểm tương đồng sẽ được đặt (nhóm lại) vào cùng một lớp và được dán nhãn tương ứng. Mục tiêu là tạo ra các cụm, trong đó sự giống nhau của các mẫu trong

một cụm vượt quá sự giống nhau của dữ liệu điểm của một cụm khác liên quan đến một biện pháp tương tự đã được chọn trước. Một trong những thuật toán phân cụm phổ biến và đơn giản nhất là k-means clustering, được Lloyd đề xuất đầu tiên. Mục tiêu chính của phương pháp phân cụm k-means là chia một tập hợp điểm dữ liệu thành các cụm riêng biệt và gán nhãn các thành viên cho phù hợp. Đối với thuật toán này, số k của cụm phải được xác định trước. Cách làm phổ biến nhất là lặp đi lặp lại để cải thiện vị trí của trung tâm cụm (centroids) bằng cách áp dụng mỗi bước sau:

Bước 1: khởi tạo centroids bằng cách chọn ngẫu nhiên (không thay thế) điểm dữ liệu  $k$   $\mathbf{c}_1, \dots, \mathbf{c}_k$ .

Bước 2: Tính toán khoảng cách Euclide (xem phương trình 2.3) giữa mỗi điểm dữ liệu  $\mathbf{x}_i$  và mỗi centroids. Gán nhãn cho mỗi  $\mathbf{x}_i$  tương ứng với các centroid gần nhất:

$$y_i = \underset{1 \leq j \leq k}{\operatorname{argmin}} \delta(\mathbf{x}_i, \mathbf{c}_j) \quad (2.5)$$

Bước 3: Tính toán lại trung tâm  $\mathbf{c}_j$  bằng cách lấy trung bình các giá trị của tất cả các điểm dữ liệu  $h$  thuộc về với cụm  $C_j$  tương ứng của chúng:

$$\mathbf{c}_j = \frac{1}{h} \sum_{l \in C_j} \mathbf{x}_l \quad (2.6)$$

Bước 4: Lặp lại các bước 2 và 3 cho đến khi đạt đến số lần lặp tối đa hoặc có không có thay đổi ở tính toán lại vị trí trung tâm ở bước 3.

Khi thuật toán này kết thúc, tất cả các mẫu được gán cho một trong các cụm  $k$  và đã nhận được một nhãn tương ứng.

### 2.3.2 Phương pháp VMS tuyến tính

Phần tử chính của SVM là một siêu kết nối có thể được định nghĩa chung là:

$$\{\mathbf{x} \in \mathbb{R}^d : f(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\beta} + \beta_0 = 0\}, \quad (2.7)$$

Khi  $\boldsymbol{\beta} \in \mathbb{R}^d$  biểu thị một vector đơn vị (i.e.,  $\|\boldsymbol{\beta}\| = 1$ ) với khoảng cách  $\beta_0 \in \mathbb{R}$  đến nguồn của không gian chứa.  $\boldsymbol{\beta}$  là một đơn vị vector đơn giản hóa các định nghĩa tiếp theo và đạt được bằng cách chia tỉ lệ  $\mathbf{x}$  và  $\beta_0$  cho phù hợp.

Nhãn  $y_i$  của mẫu  $\mathbf{x}_i$  được gán tương ứng với vị trí tương đối của nó với mặt phẳng. Đối với cây quyết định nhị phân với  $y_i \in \{1, -1\}$ , hàm quyết định được xác định là:

$$y_i = \text{sgn}(f(\mathbf{x}_i)) = \text{sgn}(\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0). \quad (2.8)$$

Hàm quyết định  $f(\mathbf{x})$  tính toán khoảng cách đã xác định của một điểm dữ liệu  $\mathbf{x}$  với mặt siêu phẳng. Do đó, nhãn được gán tùy thuộc vào mặt siêu phẳng.

### 2.3.3 Phân tách tuyến tính theo không gian đầu vào

Hiện tại, chỉ những tình huống tách tuyến tính mới được sử dụng. Do đó, tồn tại ít nhất một hàm  $f(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\beta} + \beta_0$ , với  $\forall i: y_i f(\mathbf{x}_i) > 0$ , đó là chia các điểm dữ liệu có nhãn. Từ bộ hợp lệ chức năng quyết định, thuật toán SVM tìm kiếm một hàm tạo ra lề giữa các điểm dữ liệu huấn luyện của hai lớp khác nhau.

$$\begin{aligned} & \max_{\boldsymbol{\beta}, \beta_0, \|\boldsymbol{\beta}\|=1} M, \\ \text{s.t. } & y_i(\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) \geq M, \quad i = 1, \dots, N \end{aligned} \quad (2.9)$$

$M$  biểu thị khoảng cách đã xác định từ hàm quyết định đến các mẫu của một lớp. Như đã nói ở trên,  $M$  được phóng đại, sao cho một siêu kết nối tối ưu tách biệt để có thể tìm thấy hai lớp. Ràng buộc  $\|\boldsymbol{\beta}\| = 1$  được xóa bằng cách thêm nó vào một bên hạn chế:

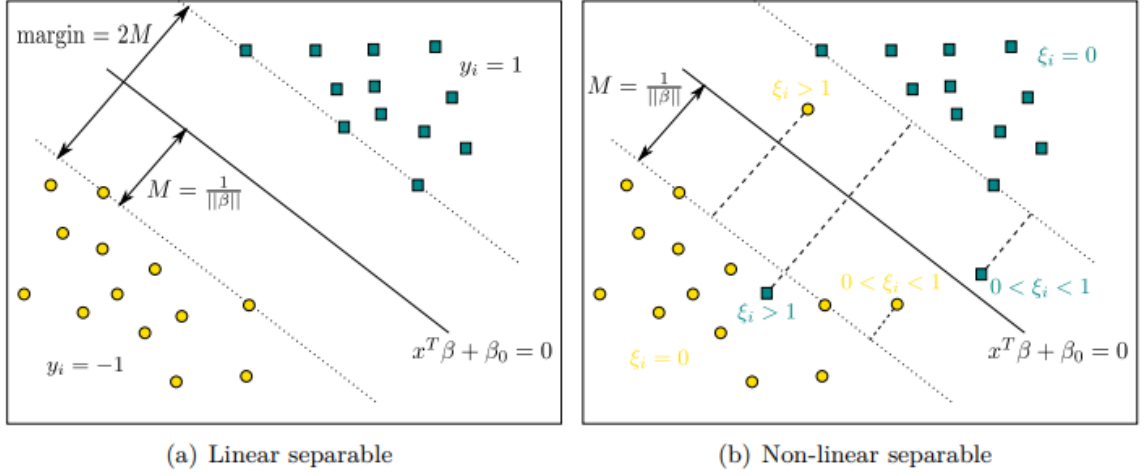
$$\begin{aligned} & \frac{1}{\|\boldsymbol{\beta}\|} y_i(\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) \geq M \\ \Leftrightarrow & y_i(\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) \geq M \|\boldsymbol{\beta}\| \\ \Leftrightarrow & y_i(\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) \geq 1 \end{aligned} \quad (2.10)$$

Bước cuối cùng có thể tắt cả  $\boldsymbol{\beta}$  và  $\beta_0$  thỏa mãn những bất bình đẳng này, bất kỳ tỷ lệ nào đáp ứng tốt là satisfying. Do đó,  $\|\boldsymbol{\beta}\|$  có thể được đặt thành  $1/M$  mà không cần các giá trị bất kỳ nào khác. Điều này dẫn đến giảm thiểu giá trị tương đương với phương trình (2.9):

$$\begin{aligned} & \min_{\boldsymbol{\beta}, \beta_0} \|\boldsymbol{\beta}\|, \\ \text{s.t. } & y_i(\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) \geq 1, \quad i = 1, \dots, N \end{aligned} \quad (2.11)$$



Bằng cách tìm một vector khả thi  $\beta$  và giá trị cho  $\beta_0$  được đảm bảo rằng tất cả các điểm dữ liệu được gán nhãn chính xác theo hàm quyết định của phương trình (2.8).



**Hình 2.9: Nguồn gốc của một máy vector hỗ trợ tuyến tính cho (a) có thể chia tách tuyến tính và (b) các điểm dữ liệu không tuyến tính.**

**Dữ liệu đầu vào có thể tách rời phi tuyến tính:** Với định nghĩa trên, chỉ có thể chia hai lớp nếu chúng bị tách rời hoàn toàn. Trong thực tế, thường có những tình huống không thể phân vùng sắc nét, tức là sự hiện diện của các ngoại lệ trong dữ liệu. Giải quyết với các loại tình huống này, cần phải chia nhỏ các ràng buộc, để các mẫu ở phía sai số cho phép. Điều này đạt được bằng cách thêm biến slack  $\xi = (\xi_1, \xi_2, \dots, \xi_N)$  đối với giới hạn biên:

$$y_i(\mathbf{x}_i^T \beta + \beta_0) \geq 1 - \xi_i, \forall i : \xi_i \geq 0, \sum_{i=1}^N \xi_i \leq c \quad (2.12)$$

Giá trị của mỗi  $\xi_i$  cho biết tổng hợp của mẫu có nhãn  $x_i$  để tham khảo với siêu kết nối:  $\xi_i = 0$  cho các mẫu ở bên phải của hàm,  $0 < \xi_i < 1$  cho các điểm dữ liệu trên lề, và  $\xi_i > 1$  cho các mẫu ở phía sai của siêu mặt phẳng. Tham số do người dùng định nghĩa  $c$  đặt một giới hạn trên cho định vị phạm. Bằng cách sử dụng các bước từ phương trình (2.10) và thêm các biến slack vào vấn đề tối ưu hóa phương trình (2.11)

$$\begin{aligned} & \min_{\beta, \beta_0} \|\beta\|, \\ & \text{s.t. } y_i(\mathbf{x}_i^T \beta + \beta_0) \geq 1 - \xi_i, \forall i, \xi_i \geq 0, \sum \xi_i \leq c. \end{aligned} \quad (2.13)$$

Tuy nhiên, để thuận tiện, sử dụng hình thức sau với vấn đề tối ưu hóa

$$\begin{aligned} \min_{\beta, \beta_0} \quad & \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i, \\ \text{s.d.} \quad & \xi_i \geq 0, \quad y_i(\mathbf{x}_i^T \beta + \beta_0) \geq 1 - \xi_i \quad \forall i \end{aligned} \quad (2.14)$$

Điều này là hợp lệ, vì nó có thể mở rộng  $\beta$  và  $\beta_0$  như mong muốn. Nên slack là một phần cố định của việc giảm thiểu và thuật ngữ

$$C \sum_{i=1}^N \xi_i$$

Là để kiểm soát cách nhiều mẫu bị thất lạc trong tham chiếu đến các lớp trên. Thông số  $C > 0$  điều chỉnh các điểm dữ liệu không đúng chỗ này:  $C$  càng được sử dụng nhiều thì sẽ ít điểm sai số hơn của siêu mặt phẳng.

Phương trình (2.14) là một hình thức cổ điển hóa của một vấn đề tối ưu hóa lồi hoặc nhiều hơn cụ thể là: một vấn đề tối ưu hóa hạn chế tuyến tính với hàm mục tiêu bậc hai. Một cách cơ bản để xử lý các vấn đề của biểu mẫu này là xác định lại hệ số sử dụng Lagrange và phải giải quyết vấn đề tối ưu hóa mới với phương trình bậc hai. Áp dụng các hệ số Lagrange cho phương trình (2.14) dẫn đến.

$$L = \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i [y_i(\mathbf{x}_i^T \beta + \beta_0) - (1 - \xi_i)] - \sum_{i=1}^N \mu_i \xi_i, \quad (2.15)$$

Trong đó  $\beta$ ,  $\beta_0$  và  $\xi_i$  được giảm thiểu trong khi các hệ số nhân Lagrange mới được sử dụng  $\alpha_i$  và  $\mu_i$  cần được tối đa hóa. Phát sinh một phần các biến số trước đó.

$$\beta = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i, \quad (2.16)$$

$$0 = \sum_{i=1}^N \alpha_i y_i, \quad (2.17)$$

$$\alpha_i = C - \mu_i, \quad \forall i, \quad (2.18)$$

với  $\alpha_i$ ,  $\mu_i$  và  $\xi_i \geq 0 \quad \forall i$ . Ở đây, sử dụng phương trình (2.14) giúp lấy được phương trình (2.15) dễ dàng hơn. Bằng cách áp dụng phương trình (2.16) và (2.18) trên phương trình (2.15) về vấn đề tối ưu hóa kép



$$\begin{aligned} \max_{0 \leq \alpha_i \leq C} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0. \end{aligned} \quad (2.19)$$

Để tìm ra giải pháp tối ưu  $\alpha_i^*$ , điều kiện Karush-Kuhn-Tucker được giới thiệu, bổ sung các điều kiện (2.16) - (2.18) vào phương trình sau:

$$\alpha_i [y_i (\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) - (1 - \xi_i)] = 0, \quad (2.20)$$

$$\mu_i \xi_i = 0, \quad (2.21)$$

$$y_i (\mathbf{x}_i^T \boldsymbol{\beta} + \beta_0) - (1 - \xi_i) \geq 0, \quad (2.22)$$

Với  $i = 1, \dots, N$ . Các giải pháp của các điều kiện Karush-Kuhn-Tucker là đủ tối ưu cho phương trình bậc hai, và hơn nữa bằng với tối ưu toàn bộ, vấn đề tối ưu hóa ban đầu chỉ là với các ràng buộc tuyến tính. Tất cả các cách  $\alpha_i^* > 0$  các mẫu  $\mathbf{x}_i$  đáp ứng điều kiện (2.20) trong khi giữ nguyên điều kiện (2.22) được gọi là vector hỗ trợ. Do đó, các vector hỗ trợ là các mẫu đào tạo được đặt trên lề hoặc ở phía đối diện của siêu phẳng (xem phương trình (2.22)). Những vector này cho phép giải quyết vấn đề ban đầu (2.14) như là một giải pháp tối ưu của  $\boldsymbol{\beta}$  được cho bởi

$$\boldsymbol{\beta}^* = \sum_{i=1}^N \alpha_i^* y_i \mathbf{x}_i \quad (2.23)$$

được lấy trực tiếp từ phương trình (2.16). Điều duy nhất còn lại để tìm là một giải pháp  $\beta_0^*$ . Đối với điều này, tất cả các mẫu đào tạo với  $\xi_i^* = 0$  cũng tuân thủ  $0 < \alpha_i^* < C$  xem các ràng buộc (2.21) và (2.18) được áp dụng. Vì vậy, đây là những điểm dữ liệu được đặt chính xác ở rìa của lề. Sau khi tìm tất cả các giải pháp, lớp của một mẫu mới  $\mathbf{x}'$  được thiết lập bởi

$$y_i = \text{sgn}[f^*(\mathbf{x}')] = \text{sgn}[\mathbf{x}'^T \boldsymbol{\beta}^* + \beta_0^*]. \quad (2.24)$$

### 2.3.4 Phương pháp VMS phi tuyến tính

Dữ liệu dạng không gian đầu vào luôn được giả định là có thể tách bởi một hàm tuyến tính, nhưng trong nhiều lĩnh vực, đây không phải là cách xử lý. Để đạt được phân loại hợp lý trong không gian phi tuyến tính, chức năng chính cần được thay đổi. SVM có thể điều chỉnh kết quả của mình bằng cách sử dụng chức năng ánh xạ đặc trưng  $\Phi$ . Bản đồ chức năng này từ không gian tách rời phi tuyến tính đến không gian được tính bằng chiều cao  $H$  nơi có thể tách tuyến tính.

Quay trở lại bài toán tối ưu hóa kép trong phương trình (2.19). Điều này giúp tích hợp bản đồ tính năng  $\Phi$  vì nó chỉ cần thay thế phần này bằng một kết quả có chứa ánh xạ tới không gian chiều cao hơn. Điều này dẫn đến vấn đề tối ưu hóa kép cho phi tuyến tính SVM:

$$\begin{aligned} \max_{0 \leq \alpha_i \leq C} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0. \end{aligned} \quad (2.25)$$

Để tránh tính toán rõ ràng của  $\Phi : \mathbb{R}^d \rightarrow H$ , ánh xạ được gây ra bởi một kernel  $K$  với  $K(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}') \rangle$ . Kỹ thuật này được gọi là kỹ thuật hạt nhân. Do đó, chỉ cần áp dụng hàm hạt nhân trên các mẫu và tính toán kết quả sau đó thay vì tính toán một kết quả trong một chiều cao hơn. Một hạn chế là ma trận hạt nhân  $K \in \mathbb{R}^{n \times n}$  được sử dụng trong lập trình bậc hai là semidefinite, có nghĩa là điều kiện  $\mathbf{x}^T K \mathbf{x} \geq 0$  phải áp dụng. Với điều kiện này, nó được đảm bảo rằng có thể tách các điểm dữ liệu trong không gian chiều cao hơn.

Ba hàm kernel phổ biến thỏa mãn ràng buộc này là:

$$\text{Linear kernel: } K_{lin}(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle \quad (2.26)$$

$$\text{Polynomial kernel: } K_{poly}(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle^p, p \in \mathbb{N} \quad (2.27)$$

$$\text{Radial basis function (RBF): } K_{rbf}(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2), \gamma \in \mathbb{R}, \gamma > 0 \quad (2.28)$$

Bằng cách này, hạt nhân tuyến tính tương đương với SVM tuyến tính được giải thích ở trên. Sự gia tăng tính phức tạp của tính toán bây giờ hoàn toàn phụ thuộc vào độ phức tạp của hàm hạt nhân và kích thước dữ liệu của không gian đầu vào

ban đầu. Chi tiết hơn thông tin về các hàm hạt nhân có thể được tìm thấy trong Hastie et al.

Thêm hàm hạt nhân vào bài toán tối ưu hóa kép của phương trình (2.25) trong

$$\begin{aligned} \max_{0 \leq \alpha_i \leq C} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0. \end{aligned} \quad (2.29)$$

và nhãn của mẫu mới  $\mathbf{x}'$  được xác định với hàm quyết định liên quan là sau:

$$\begin{aligned} y_i &= \text{sgn}[f^*(\mathbf{x}')] = \text{sgn}[\Phi(\mathbf{x}')^T \boldsymbol{\beta}^* + \beta_0^*] \\ &= \text{sgn}\left[\sum_{i=1}^N \alpha_i^* y_i \langle \Phi(\mathbf{x}'), \Phi(\mathbf{x}_i) \rangle + \beta_0^*\right] \\ &= \text{sgn}\left[\sum_{i=1}^N \alpha_i^* y_i \mathcal{K}(\mathbf{x}', \mathbf{x}_i) + \beta_0^*\right]. \end{aligned} \quad (2.30)$$

### 2.3.5 Đánh giá và phân tích kết quả

Ở phần này, các thông tin cơ bản về học máy, bao gồm cả tự động việc học các hàm mà ánh xạ các giá trị từ đầu vào đến không gian đầu ra, được giới thiệu. Giá trị của không gian đầu vào được gọi là các mẫu và các đối lập của chúng trong không gian đầu ra là các nhãn. Trong suốt công việc này, các đặc điểm của thuật ngữ được sử dụng để mô tả các thành phần đơn lẻ của một mẫu. Để huấn luyện, kiểm tra và đánh giá các mô hình phân loại và hồi quy, cũng có các bộ dữ liệu tương ứng được giới thiệu. Với kiến thức về đào tạo các tập dữ liệu, một sự phân biệt giữa việc học có giám sát và không giám sát là có thể. Phương pháp được giám sát sử dụng các mối liên hệ đã biết giữa các mẫu và nhãn của một khóa đào tạo tập dữ liệu để tìm hiểu các hàm ánh xạ. Các chức năng này sau này được sử dụng để chỉ định thích hợp nhãn cho các mẫu mới. Để đưa ra một ví dụ đơn giản về mô hình phân loại và hồi quy, thuật toán học được giám sát k-láng giềng gần nhất được sử dụng. Vì chủ yếu sử dụng phương pháp học được giám sát, chỉ có một ví dụ ngắn về một phương pháp học không giám sát, phương pháp được đưa ra là: k-means clustering. Với cách phân cụm k, có thể phân loại các mẫu thành các nhóm k (cụm) và gán các nhãn tương ứng cho chúng. Trên cơ sở đó, phần thứ hai của chương này

tập trung vào việc nghiên cứu các nguyên tắc sử dụng cơ bản của học máy vector hỗ trợ (SVM). Bắt đầu từ tối ưu hóa mô hình SVM tuyến tính cho dữ liệu dạng không gian đầu vào có thể phân tách tuyến tính. Dựa trên điều này, việc phân loại phi tuyến tính không gian bằng cách áp dụng các biến slack được sử dụng. Kết luận được gọi là các hàm hạt nhân cho mô hình, vấn đề tối ưu hóa được tham chiếu vào một không gian đầu vào chiều cao hơn, trong đó có thể tách tuyến tính. Mô hình SVM có thể phân loại các mẫu không gian dữ liệu thậm chí phức tạp và gán các nhãn thích hợp, tuân thủ hai tham số do người dùng định nghĩa  $C$  và  $\gamma$ , điều chỉnh độ phức tạp của học máy. Phần cuối cùng của chương này giải thích những thay đổi cần thiết để tạo ra kết quả hồi quy bằng cách sử dụng SVM. Thay vì các giá trị rời rạc, các không gian dữ liệu có giá trị liên tục. Phần mở rộng của SVM mô hình hồi quy được gọi là hỗ trợ hồi quy vector (SVR). Trong luận văn này, SVR được sử dụng làm mô hình chính dự báo nhu cầu sử dụng điện mặt trời.

## **2.4 Kết luận chương 2**

Trong chương này, luận văn đã trình bày tổng quan các khái niệm cơ bản liên quan tới dự báo, dự báo kinh tế. Các phương pháp thường sử dụng trong dự báo kinh tế bao gồm các phương pháp liên quan tới kinh tế lượng và học máy.

## CHƯƠNG 3: XÂY DỰNG MÔ HÌNH DỰ BÁO TÌNH HÌNH SỬ DỤNG ĐIỆN MẶT TRỜI

### 3.1 Môi trường cài đặt

Chương trình được viết trên nền.NET Framework sẽ được triển khai trong môi trường phần mềm. Môi trường phần mềm này là một máy ảo trong đó cung cấp các dịch vụ như an ninh phần mềm (*security*), quản lý bộ nhớ (*memory management*), và các xử lý lỗi ngoại lệ (*exception handling*).

.NET framework bao gồm tập các thư viện lập trình lớn, và những thư viện này hỗ trợ việc xây dựng các chương trình phần mềm như lập trình giao diện; truy cập, kết nối cơ sở dữ liệu; ứng dụng web; các giải thuật, cấu trúc dữ liệu; giao tiếp mạng... CLR cùng với bộ thư viện này là 2 thành phần chính của.NET framework.

.NET framework đơn giản hóa việc viết ứng dụng bằng cách cung cấp nhiều thành phần được thiết kế sẵn, người lập trình chỉ cần học cách sử dụng và tùy theo sự sáng tạo mà gắn kết các thành phần đó lại với nhau. Nhiều công cụ được tạo ra để hỗ trợ xây dựng ứng dụng.NET, và IDE (*Integrated Development Environment*) được phát triển và hỗ trợ bởi chính Microsoft là Visual Studio.

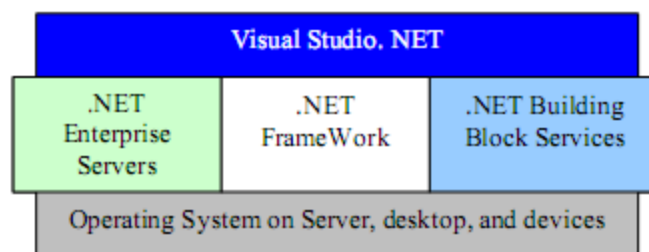
*Các thức cài đặt chương trình.*

Chương trình chạy trên nền windows.

### 3.2 Ngôn ngữ sử dụng

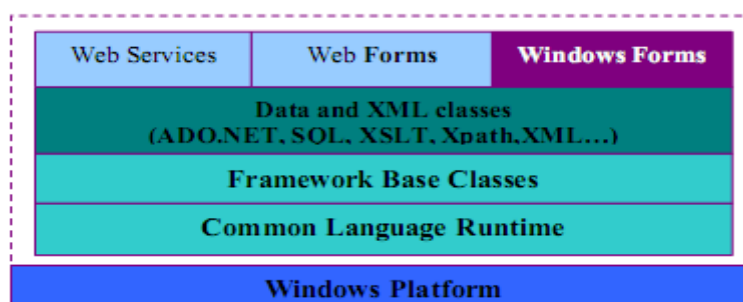
#### Microsoft .Net

*Kiến trúc của Microsoft .NET*



**Hình 3.1: Thành phần chính của Microsoft .NET**

## Microsoft .NET Framework



**Hình 3.2: Microsoft .NET Framework**

### Ngôn ngữ lập trình C#

Ngôn ngữ C# khá đơn giản với khoảng 80 từ khóa và khoảng hơn mười kiểu dữ liệu xây dựng sẵn. Tuy nhiên, ngôn ngữ C# có ý nghĩa cao khi nó thực thi những khái niệm lập trình hiện đại. C# bao gồm tất cả những hỗ trợ cho cấu trúc, thành phần, lập trình hướng đối tượng. Những tính chất đó hiện diện trong ngôn ngữ lập trình hiện đại và ngôn ngữ C# hội tụ đủ những điều kiện như vậy, hơn nữa nó được xây dựng trên nền tảng của hai ngôn ngữ mạnh nhất là C++ và java.

### 3.3 Kiến trúc hệ thống

Dựa vào việc xác định các biến tham gia vào mô hình dự báo cũng như phương pháp thu thập dữ liệu và các nguồn dữ liệu đã trình bày ở trên. Mô hình sử dụng bộ số liệu từ các biến độc lập tại chu kỳ thứ  $n$  để dự báo cho biến phụ thuộc tại chu kỳ thứ  $n+1$ . Mô hình không chỉ có ý nghĩa dự báo cho kết quả của biến phụ thuộc (nhu cầu sử dụng năng lượng mặt trời) ở chu kỳ kế tiếp mà còn thể hiện mức độ ảnh hưởng của các biến độc lập đến biến phụ thuộc nhằm mục tiêu kiểm soát các biến độc lập để đảm tính ổn định của nhu cầu sử dụng năng lượng mặt trời.

Các dữ liệu cần đưa vào là dữ liệu của các biến độc lập còn lại sau khi đã loại các biến không thể hiện được sự tương quan. Dựa vào dữ liệu của chu kỳ thứ  $n$  để dự báo cho chu kỳ thứ  $n+1$ , các hệ số trong mô hình sẽ thay đổi theo thời gian vì vậy cần phải liên tục tính toán lại tính phù hợp của mô hình theo thời gian để tránh sai lệch.

#### *Yêu cầu của hệ thống*

Hệ thống dự báo nhu cầu sử dụng bao gồm 2 pha chính: pha huấn luyện và pha dự báo.

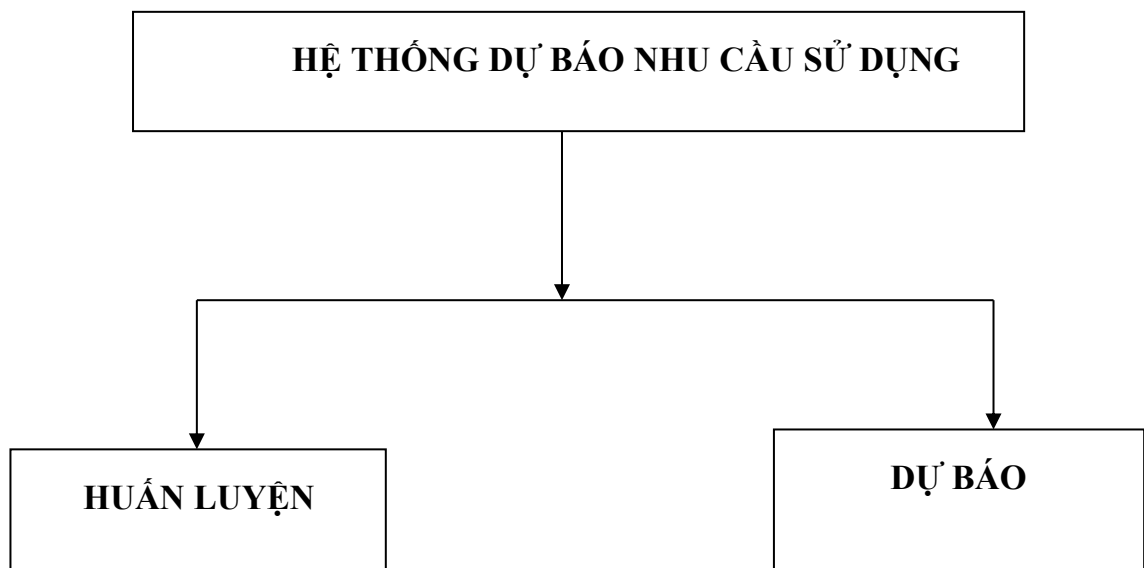
- Pha huấn luyện: pha huấn luyện thực hiện việc huấn luyện dữ liệu dựa trên mạng nơ ron nhân tạo.

- Pha dự báo: pha dự báo sử dụng bộ dữ liệu với các thông số đầu vào đảm bảo với yêu cầu, cùng kiểu định dạng dữ liệu và thực hiện cho kết quả đầu ra.

### ***Các chức năng cơ bản của hệ thống***

Chức năng tổng quát

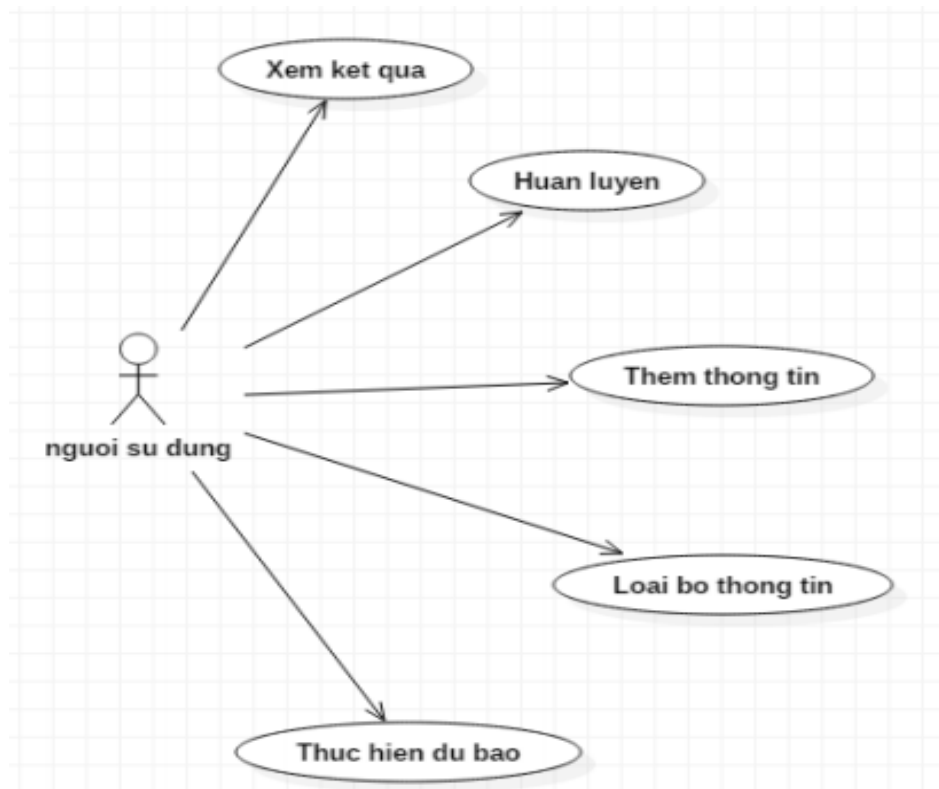
Dựa vào yêu cầu của hệ thống dự báo giá điện năng, chức năng chính của hệ thống được mô tả như hình vẽ 3.8 dưới đây



***Hình 3.3: Sơ đồ chức năng hệ thống dự báo***

Hình 3.4 dưới đây mô tả biểu đồ ca sử dụng tổng quát của hệ thống trong biểu đồ use case tổng quát.

- o Huấn luyện
- o Thêm dữ liệu
- o Xem thông tin
- o Dự báo



**Hình 3.4: Biểu đồ Use case tổng quát**

### **Biểu đồ lớp**

Sử dụng Office Excel để tạo cơ sở dữ liệu lưu trữ quá trình huấn luyện và kết quả dự báo nhu cầu sử dụng cho lần tiếp theo. Chương trình được thiết kế từ các bảng sau:

**Bảng 3.1: Dữ liệu huấn luyện**

| Tên trường | Kiểu dữ liệu | Khóa | Mô tả                       |
|------------|--------------|------|-----------------------------|
| ID         | Int          | X    |                             |
| SHCS       | Int          |      | Số giờ chiếu sáng           |
| SLPhat     | Double       |      | Sả lượng phát điện          |
| SLKH       | Int          |      | Số lượng khách hàng sử dụng |
| TGTK       | Datetime     |      | Thời gian thống kê          |



**Bảng 3.2: Lưu dự báo qua các lần thử nghiệm**

| Tên trường    | Kiểu dữ liệu   | Khóa | Mô tả                                        |
|---------------|----------------|------|----------------------------------------------|
| ID            | Int            | X    |                                              |
| NhuCauDuBao   | decimal(18, 0) |      | Nhu cầu sử dụng năng lượng dự báo            |
| ThoiGianDuBao | Datetime       |      | Thời gian dự báo thay đổi nhu cầu năng lượng |
| NgayTao       | datetime       |      | Ngày tạo                                     |
| LanDuBao      | Int            |      | Lần dự báo                                   |
| SL_Truoc      | Int            |      | Số liệu gần nhất trước khi dự báo            |
| TyLeChenh     | decimal(18, 0) |      | Tỷ lệ chênh so với thực tế                   |

### 3.4 Quy trình thực nghiệm

Tiến hành thực nghiệm trên hai tập dữ liệu: dữ liệu thực và dữ liệu sinh tự động.

#### Sử dụng phương thức k - fold

Mỗi tập dữ liệu, sử dụng phương thức k – fold, chia tập dữ liệu làm 10 fold, sau đó tiến hành huấn luyện với 10 lần lặp. Mỗi lần, sử dụng 9 – fold dữ liệu làm tập huấn luyện mô hình, fold còn lại làm tập test. Mô hình thực nghiệm luận văn mới chỉ thực hiện thực nghiệm với mô hình SVR đơn. Do điều kiện về thời gian nên việc thực hiện thực nghiệm với mô hình SVR có kết hợp giải thuật di truyền chưa được hoàn thiện. Vì vậy trong luận văn này, chưa trình bày phần thực nghiệm với mô hình SVR có kết hợp giải thuật di truyền.

Lựa chọn 2 hàm nhân chính cho SVR là nhân Radial Basis Functions (RBF) và nhân đa thức (Polynomial - Poly) để ánh xạ dữ liệu như là các biến đầu vào xi trong mô hình SVR. Thử nghiệm thực hiện với giải thuật smoreg cho mô hình hồi quy vector hỗ trợ.

Đầu vào của mô hình bao gồm: số lượng người lắp đặt thực tế, sản lượng phát điện đã đăng ký, tổng số giờ chiếu sáng theo tháng trong năm. Đầu ra của mô hình là số lượng người đăng ký lắp đặt.

#### Bảng 3.3: Lựa chọn giá trị các tham số cho mô hình thực nghiệm.

| Tham số                       |            | Dữ liệu thực | Dữ liệu sinh tự động |
|-------------------------------|------------|--------------|----------------------|
| Hàm nhân đa thức<br>(bậc = 1) | C          | 2            | 2                    |
|                               | $\epsilon$ | 0.0001       | 0.0001               |
| Hàm nhân RBF<br>(Gama = 0.01) | C          | 2            | 1                    |
|                               | $\epsilon$ | 0.0001       | 0.0001               |

### Phương pháp đánh giá mô hình

Trong nghiên cứu này, sử dụng các phương pháp sau

$$- RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2} \text{ (Sai số căn quân phương)}$$

$$- R = \frac{\sum_{i=1}^n (y_i - \bar{y})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (f_i - \bar{f})^2}} \text{ (Sai số tương quan)}$$

$$- MAE = \frac{1}{n} \sum_{i=1}^n |y_i - f_i| \text{ (Sai số tuyệt đối trung bình)}$$

Trong đó,  $y_i$  và  $f_i$  chỉ giá trị thực và giá trị dự báo tại thời điểm  $i$ ,  $\bar{y}$ ,  $\bar{f}$  là giá trị trung bình của giá trị thực và giá trị dự báo.

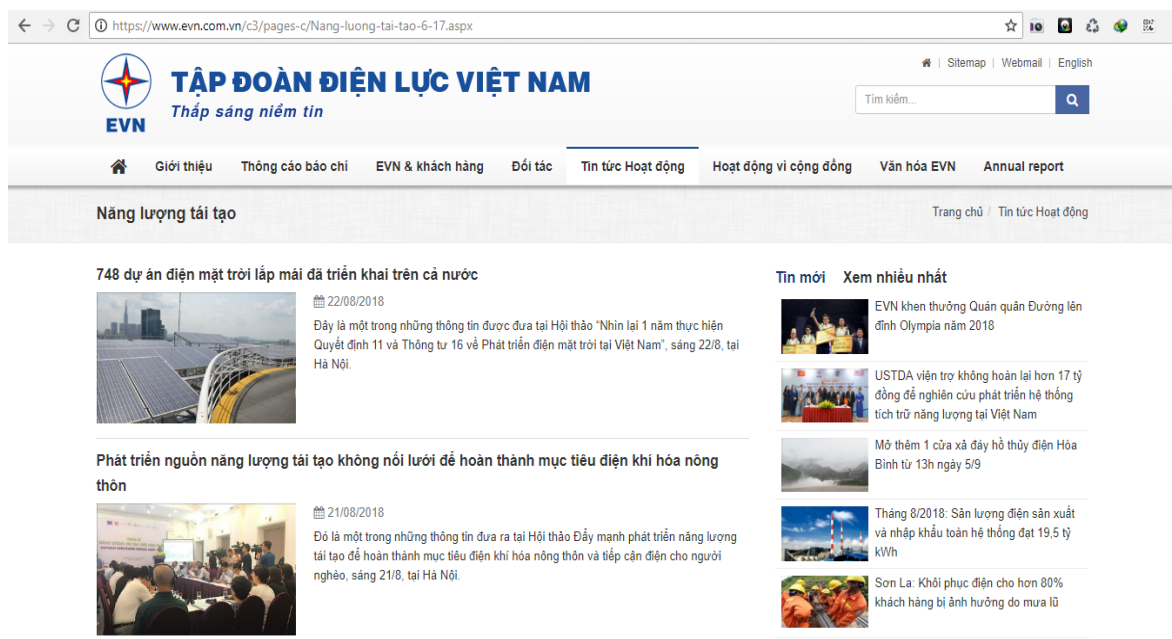
Phương pháp dự báo tốt là phương pháp nhận được sai số  $R$  lớn còn sai số  $RMSE$  và  $MAE$  nhỏ (càng gần về không càng tốt).

## 3.5 Kết quả thử nghiệm

### 3.5.1 Nguồn dữ liệu sử dụng

Dữ liệu dùng cho thử nghiệm được lấy từ 02 nguồn:

- Dữ liệu thống kê số lượng khách hàng sử dụng và công suất phát của pin mặt trời được lấy từ Ban kinh doanh điện năng trực thuộc Tập đoàn Điện lực Việt Nam.



**Hình 3.5: Tập đoàn Điện lực Việt Nam**

- Dữ liệu số giờ chiếu sáng trong tháng được lấy từ tổng cục thống kê Việt Nam.

**Hình 3.6: Tổng cục thống kê Việt Nam**

### 3.5.2 Dữ liệu dùng cho thử nghiệm

Với 03 nhân tố ảnh hưởng, luận văn thực hiện thử nghiệm với các dữ liệu từ tháng 02/2018 đến tháng 07/2018 như sau:

**Bảng 3.4: Dữ liệu thống kê về khách hàng và sản lượng phát**

| STT | Công ty Điện lực | Luỹ kế đến tháng 2  |                      | Tháng 3             |                      | Tháng 4             |                      | Tháng 5             |                      | Tháng 6             |                      | Tháng 7             |                      |
|-----|------------------|---------------------|----------------------|---------------------|----------------------|---------------------|----------------------|---------------------|----------------------|---------------------|----------------------|---------------------|----------------------|
|     |                  | Số lượng khách hàng | Sản lượng phát (kWh) | Số lượng khách hàng | Sản lượng phát (kWh) | Số lượng khách hàng | Sản lượng phát (kWh) | Số lượng khách hàng | Sản lượng phát (kWh) | Số lượng khách hàng | Sản lượng phát (kWh) | Số lượng khách hàng | Sản lượng phát (kWh) |
| 1   | An Giang         |                     |                      |                     |                      |                     |                      | 1                   | 133                  | 1                   | 335                  | 4                   | 760                  |
| 2   | Bà Rịa Vũng Tàu  |                     |                      |                     |                      |                     |                      | 2                   | -                    | 7                   | -                    | 35                  | 4.647                |
| 3   | Bạc Liêu         |                     |                      | 1                   | -                    | 3                   | 3.240                | 3                   | 3.601                | -                   | 2.216                | -                   | 2.587                |
| 4   | Bến Tre          |                     |                      | 1                   | 2.627                | 1                   | 330                  | 1                   | 500                  | 1                   | 647                  | 3                   | 538                  |
| 5   | Bình Dương       | -                   | -                    | 2                   | 5.230                | 3                   | 2.920                | 4                   | 7.221                | 12                  | 6.474                | 19                  | 10.355               |
| 6   | Bình Phước       |                     |                      |                     |                      |                     |                      |                     |                      | 1                   | 513                  | 2                   | 522                  |
| 7   | Bình Thuận       |                     |                      |                     |                      |                     |                      |                     |                      | 7                   | 64                   | 8                   | 964                  |
| 8   | Cà Mau           | 1                   | 4.000                | 1                   | 5.200                | 1                   | 9.170                | 1                   | 8.360                | 1                   | 6.730                | 1                   | 6.830                |
| 9   | Cần Thơ          |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      |
| 10  | Đồng Nai         |                     |                      |                     |                      |                     |                      |                     |                      | 2                   | 506                  | 5                   | 8.366                |
| 11  | Đồng Tháp        | 2                   | 87.106               | 5                   | 4.323                | 6                   | 3.945                | 8                   | 3.445                | 10                  | 44.798               | 13                  | 14.043               |
| 12  | Hậu Giang        |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      | 1                   | -                    |
| 13  | Kiên Giang       |                     |                      | 1                   | 26                   | 1                   | 993                  | 1                   | 2.398                | 6                   | 7.654                | 7                   | 2.398                |
| 14  | Lâm Đồng         |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      |
| 15  | Long An          |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      | 7                   | 515                  |
| 16  | Ninh Thuận       | -                   | -                    | -                   | -                    | -                   | -                    | -                   | -                    | 8                   | -                    | 4                   | 2.577                |
| 17  | Tây Ninh         |                     |                      |                     |                      |                     |                      |                     |                      | 5                   | 744                  | 5                   | 2.013                |
| 18  | Tiền Giang       |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      |                     |                      |
| 19  | Trà Vinh         | -                   | -                    | -                   | -                    | -                   | -                    | 1                   | 26                   | 1                   | 55                   | 3                   | 376                  |
| 20  | Sóc Trăng        | 4                   | 2.887                | 4                   | 3.413                | 4                   | 3.381                | 4                   | 2.637                | 6                   | 2.055                | 7                   | 3.041                |
| 21  | Vĩnh Long        |                     |                      |                     |                      | 3                   | 15.302               | 3                   | 18.433               | 3                   | 16.449               | 4                   | 14.197               |
|     | Tổng cộng        | 7                   | 93.993               | 15                  | 20.819               | 22                  | 39.281               | 29                  | 46.754               | 71                  | 89.240               | 128                 | 74.729               |

### 3.6 Một số giao diện của chương trình

 DỰ BÁO NHU CẦU SỬ DỤNG NĂNG LƯỢNG MẶT TRỜI

NHẬP DỮ LIỆU HUẤN LUYỆN

|   | Thời điểm thống kê | Số lượng khách hàng | Sản lượng phát (kWh) | Số giờ chiếu sáng | Tỉ lệ |
|---|--------------------|---------------------|----------------------|-------------------|-------|
| ▶ | 2/2018             | 7                   | 93993                | 144,3             | 0     |
|   | 3/2018             | 15                  | 20819                | 158,5             | 8     |
|   | 4/2018             | 22                  | 39281                | 159,8             | 7     |
|   | 5/2018             | 29                  | 46754                | 168,9             | 7     |
|   | 6/2018             | 71                  | 89240                | 210,5             | 42    |
| * |                    |                     |                      |                   |       |

*Hình 3.7: Giao diện dữ liệu huấn luyện cho chương trình*

NHẬP SỐ LIỆU DỰ BÁO

Sản lượng phát (kWh)

Số giờ chiếu sáng  Chọn tháng  Chọn năm

Số lượng khách hàng dự báo: 125

Số lượng khách hàng thực tế:

Tỉ lệ: 54

Sai số: 2,3715 %

***Hình 3.8: Giao diện dự báo nhu cầu sử dụng điện mặt trời***

### **3.7 Kết luận chương 3**

Trong chương này, luận văn đã trình bày quá trình xây dựng, cài đặt hệ thống và thử nghiệm kết quả trên hệ thống. Đồng thời chương này cũng đánh giá lỗi tỉ lệ hay còn gọi là sai số chênh lệch giữa kết quả dự báo và kết quả thực. Các giao diện của hệ thống và mô tả hệ thống cũng đã được thực hiện tại chương này.

## KẾT LUẬN

### Kết quả đạt được

- Từ mô hình máy vecto hỗ trợ đã xây dựng được tiến hành dự báo nhu cầu sử dụng năng lượng mặt trời năm 2018 từ tháng 7/2018, từ kết quả đó so sánh với số liệu thống kê đã thu thập để kiểm định kết quả của mô hình dự báo với phương sai hiệu chỉnh là 2,371541502.

### Đánh giá ưu nhược điểm

#### Ưu điểm.

- Giúp người dự báo biết những biến nào có ảnh hưởng nhiều đến biến phụ thuộc. Điều này có ý nghĩa quan trọng vì dựa trên cơ sở đó để xếp thứ tự ưu tiên, người sử dụng sẽ có các chiến lược thích hợp.

- Từ việc đánh giá kết quả sau khi dự báo với số liệu trong quá khứ nhận thấy mô hình dự báo tương đối chính xác trong vòng nhiều tháng nên rất thích hợp cho việc dự báo dài hạn để đưa ra các định hướng kinh tế trong tương lai.

#### Nhược điểm

Việc dự báo bằng mô hình còn gặp phải các nhược điểm:

- Không tiên liệu trước được độ trễ.
- Trong quá trình dự báo còn đưa ra nhiều giả định
- Mô hình có thêm một độ trễ thì mất đi một bậc tự do, nếu dữ liệu chuỗi thời gian không đủ dài thì ý nghĩa thống kê của mô hình ngày càng kém.
- Các biến giải thích thực chất là giá trị của một biến X theo thời gian, điều này gây ra sự tương quan giữa các biến giải thích trong mô hình, tức là có hiện tượng đa cộng tuyến. Ước lượng các tham số của mô hình trong trường hợp có đa cộng tuyến sẽ cho kết quả kém chính xác.

### Hướng phát triển trong tương lai

Có thể nói mô hình trên có tính thực tế và ứng dụng thực tiễn rất cao, nhất là trong thời đại kinh tế đang phát triển nhanh chóng như hiện nay. Vậy nên cần sử dụng mô hình trên như một tiền đề để tiếp tục nghiên cứu và phát triển. Cần có cái nhìn tổng quan và đánh giá chính xác hơn năng lực dự báo của mô hình từ đó đề ra

các biện pháp khắc phục nhược điểm, đề ra các hướng phát triển cho mô hình trong tương lai, áp dụng để dự báo chỉ số giá tiêu dùng cho các năm sắp tới.

Ngoài ra, có thể dựa trên mô hình để triển khai xây dựng thành một phần mềm ứng dụng, giúp việc dự báo trở nên chuyên nghiệp, thuận tiện, phổ biến trên quy mô rộng và hạn chế thấp nhất sai sót có thể xảy ra ảnh hưởng đến kết quả dự báo. Giúp đưa mô hình dự báo chỉ số giá tiêu dùng bằng hồi quy bội trở thành công cụ đắc lực hỗ trợ việc dự báo, phát huy hết vai trò và tác dụng của dự báo.

Không dừng lại ở đó, cần tiếp tục nghiên cứu để mở rộng lĩnh vực dự báo, đặc biệt là các lĩnh vực có nhiều biến động, tốc độ phát triển nhanh chóng và ảnh hưởng nhiều đến nền kinh tế.

## TÀI LIỆU THAM KHẢO

- [1] Solar electric power association, Miriam Makhyoun 2011.
- [2] H. Madsen: Time Series Analysis, Chapman and Hall, 392 pp, 2008
- [3] J.M. Morales, A.J. Conejo, H. Madsen, P. Pinson, M. Zugno: *Integrating Renewables in Electricity Markets*, Springer, 430 pp., 2013
- [4] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, C. Draxl: *The state-of-the-art in short-term prediction of wind power*, ANEMOS plus report, 2011.
- [5] P. Pinson, H. Madsen: *Adaptive modeling and forecasting of wind power fluctuations with Markov-switching autoregressive models*. Journal of Forecasting, 2010.
- [6] Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften: *Support Vector Regression for Solar Power Prediction*. Oldenburg, den 12. April 2017
- [7] Tổng cục thống kê, 2008, “*Niên giám Thống Kê*”, NXB. Thống kê.
- [8] Tập đoàn Điện lực Việt Nam  
Nguồn: [www.evn.com.vn](http://www.evn.com.vn)
- [9] Tổng cục Thống kê Việt Nam  
Nguồn: [www.gso.gov.vn](http://www.gso.gov.vn)