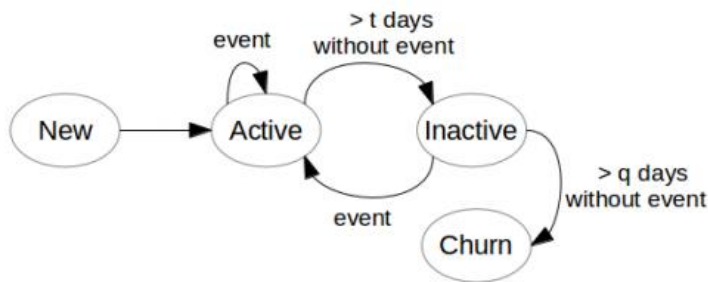


## 1. Định nghĩa thuê bao ngủ đông

- Thuê bao ngủ đông (inactive user): là một hiện tượng khi mà hoạt động tiêu dùng của người dùng ở mức thấp không đáng kể hoặc không tiêu dùng xét trong một khoảng thời gian đủ dài nào đó. Tuy nhiên thuê bao đó có thể tiêu dùng bình thường trở lại ở thời gian sau đó.

- Trong đó, một thuê bao rời mạng là thuê bao đó sẽ dừng hẳn mọi hoạt động tiêu dùng, rời khỏi mạng không bao giờ quay trở lại nữa.

Hình dưới đây mô tả vòng đời hoạt động của một thuê bao (tham khảo [1]):



Một thuê bao mới tham gia mạng sẽ được gán trạng thái là active, và nó vẫn sẽ là active nếu vẫn có hoạt động tiêu dùng thường xuyên. Tuy nhiên, khi trong 1 khoảng thời gian  $t$  ngày mà không có bất kỳ một hoạt động tiêu dùng nào hoặc chỉ tiêu dùng ở mức thấp không đáng kể thì thuê bao đó sẽ bị coi như là inactive. Và sau đó  $q$  ngày vẫn thấy thuê bao đó không có hoạt động tiêu dùng gì cả thì nó sẽ được coi như là rời mạng. Tuy nhiên trong khoảng thời gian đó nếu thuê bao đó lại có hoạt động tiêu dùng thì nó lại trở về trạng thái active. Thông thường  $t$  sẽ được xem xét là khoảng nửa tháng hoặc 1 tháng. Còn  $q$  là  $> 5$  tháng.

**Định nghĩa:** Một thuê bao được coi như là ngủ đông trong khoảng thời gian  $t$  ( $= 2$  tuần chẳng hạn) đó là nó chỉ có số ngày tiêu dùng nhiều nhất là 5 ngày và tổng mức độ tiêu dùng là không đáng kể (ví dụ chỉ dùng khoảng 1000-2000 VNĐ).

## 2. Tại sao dữ đoán thuê bao ngủ đông lại có ý nghĩa.

- Bài toán dự đoán thuê bao ngủ đông là bài toán dự đoán theo chuỗi thời gian (dự đoán liên tiếp trong từng khoảng thời gian  $t$ ) trong khi đó bài toán dự đoán thuê bao rời mạng nói một cách thô sơ thì là dự đoán một lần rời xong.

- Có thể nói trạng thái ngủ đông thường xảy ra trước trạng thái rời mạng bởi vì khoảng thời gian đầu của rời mạng ta có thể xem như là hoạt động ngủ đông. Nếu quan sát trong khoảng 5 tháng liên mà thấy thuê bao đó ngủ đông thì khi đó ta có thể kết luận thuê bao đó rời mạng => dự đoán thuê bao ngủ đông thực ra cũng có thể thực hiện được dự đoán thuê bao rời mạng.

- Trong thực tế hiện nay, hoạt động rời mạng sẽ không phổ biến bằng hoạt động ngủ đông vì Viettel là một tập đoàn viễn thông hàng đầu Việt Nam, nên số lượng thuê bao mà tự động rời mạng Viettel sẽ không đáng kể. Một điều nữa, với sự bùng nổ mạnh mẽ của mạng xã hội hiện tại, thì việc tiêu dùng của người dùng sẽ bị ảnh hưởng. Vì vậy, tình trạng ngủ đông thường hay xảy ra hơn.

=> Phán đoán ra các thuê bao ngủ đông để đưa ra chính sách phù hợp sẽ là sự chăm sóc đến một bộ phận lớn khách hàng.

### 3. Ý tưởng giải quyết bài toán.

Để cho đơn giản trong nếu ta xét khoảng thời gian xét ngủ đông là  $t = 1$  tuần chẳng hạn.

Giả sử ở thời điểm hiện tại ta quan sát được  $N$  tuần tiêu dùng của  $M$  thuê bao. Mục tiêu của ta là dự đoán xem thuê bao đó là active hay inactive trong tuần tiếp theo (tuần  $N+1$ ). Ta sẽ đưa bài toán dự đoán này về bài toán phân loại (active hay inactive).

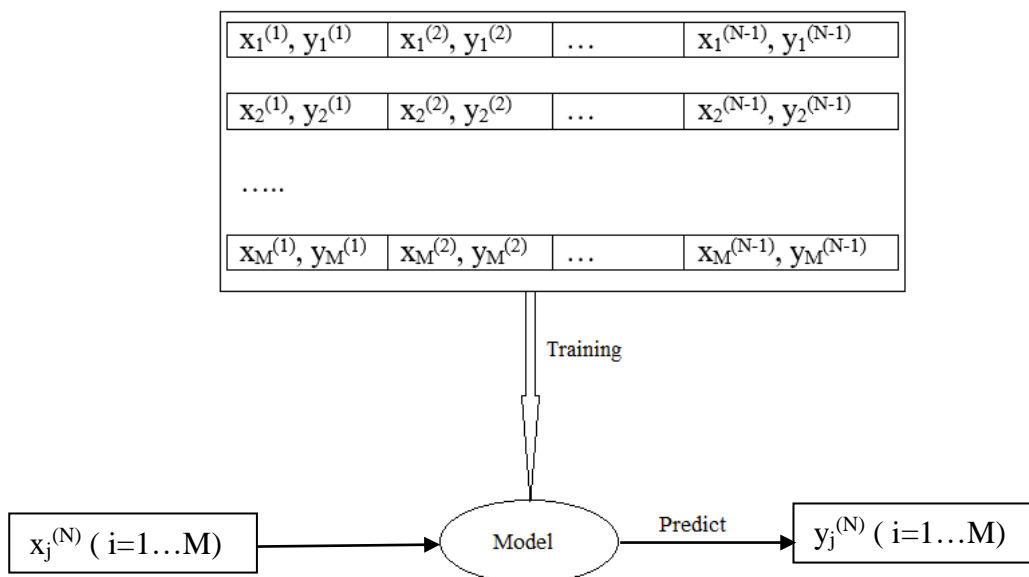
Gán nhãn cho từng tuần: Nhãn cho tuần thứ  $i$  là trạng thái hoạt động của thuê bao trong tuần thứ  $i+1$ .

Tập dữ liệu training sẽ là:  $D = \{(x_j^{(i)}, y_j^{(i)}), i = 1, 2, \dots, N-1 \text{ và } j = 1, 2, \dots, M\}$  ( $i$  là id tuần,  $j$  là id của thuê bao).

Trong đó:

- $x_j^{(i)}$  là vector đặc trưng cho hoạt động tiêu dùng của người dùng  $j$  trong tuần thứ  $i$ .
- $y_j^{(i)} \in \{-1, 1\}$  là biểu thị trạng thái inactive hay active của thuê bao  $j$  trong tuần thứ  $i+1$ .

Và sau khi training model, dữ liệu tuần thứ  $N$  được đưa vào để đưa ra dự đoán cho tuần  $N+1$ . Cứ như vậy, một tuần sau ta lại quan sát được dữ liệu tuần  $N+1$ , lúc đó lại đưa ra dự đoán cho tuần  $N+2$ . (sau 1 vài tuần dự đoán, có thể ta phải re-train lại model hoặc có thể dùng các phương pháp học streaming để update model theo dòng thời gian).



Chi tiết model: Tham khảo [2]

[1] [http://www.wiseathena.com/pdf/wa\\_dl.pdf](http://www.wiseathena.com/pdf/wa_dl.pdf)

[2]

<https://drive.google.com/file/d/1FI8EW5276SKjwjvwZ5U6Gn5fzgFerN9E/view?usp=sharing>