Approach:

1) First, it was important to gather information about the data as much as possible before making positive and negative examples.
   Feature Engineering:
   1) My approach was to basically group by the data as per Patient-Uid. Grouping would have definitely reduced the information so we needed to build some new features that contain the fluctuation. such as total sums/ division of two features and etc
   2) So I went on extracting as much as possible information from the Incident section like finding the total frequency of Incidents, totally different types of drugs/symptoms/test types per patient, and the total number of months patient has been medicated and grouped it as per Patient-Uid
   3) From the above features, I came to know that most of the patients who at least had one target drug had fewer symptoms and more number total drugs
   4) After that to avoid data loss I performed some division between features like finding incidents per month, drugs per incident, and so on
   5) The only strategy I can come with is clustering for labeling data as I thought patients who have got at least once in their life are somewhat similar in terms of symptoms per-incident and some features so I clustered the data and labeled it
   6) After clustering, I took the positive as the cluster having a higher number of patients who took TARGET drug at least once in lifetime
   7) Drew pair plot and saw the distribution of the features can come up with tree ML models as best fit