Proposal on Machine Learning Classification of dermatological diseases

Aakash Ghosh
19MS129

1st April,2023

Timeline

We are going to look at:

1. Current state
2. Problem Statement
3. Advantages, Issues and Prospective Remedies
4. A Naive Roadmap
5. Questions

Current state of affairs

Current state of affairs

We look at the current state of affairs regarding digital health records in India

Current state of affairs

We look at the current state of affairs regarding digital health records in India

1. Post diagnosis treatment and prescription is done on a handwritten basis in India.

Current state of affairs

We look at the current state of affairs regarding digital health records in India

1. Post diagnosis treatment and prescription is done on a handwritten basis in India.
2. There is no standardised digital system for storing data and using it for statistical analysis in India.

Current state of affairs

We look at the current state of affairs regarding digital health records in India

1. Post diagnosis treatment and prescription is done on a handwritten basis in India.
2. There is no standardised digital system for storing data and using it for statistical analysis in India.
3. National Health Stack and National Digital Health Mission (NDHM) are initiatives to create a digital health ecosystem in India.

Current state of affairs

We look at the current state of affairs regarding digital health records in India

1. Post diagnosis treatment and prescription is done on a handwritten basis in India.
2. There is no standardised digital system for storing data and using it for statistical analysis in India.
3. National Health Stack and National Digital Health Mission (NDHM) are initiatives to create a digital health ecosystem in India.
4. The benefits of a fully digital healthcare system include improved patient outcomes, reduced medical errors, and better management of chronic diseases.

Problem statement

Problem statement

Ideally we would like to have a system:

1. Which stores a patients' health Information

Problem statement

Ideally we would like to have a system:

1. Which stores a patients' health Information
2. Gives predictive diagnosis based on history and symptoms

Problem statement

Ideally we would like to have a system:

1. Which stores a patients' health Information
2. Gives predictive diagnosis based on history and symptoms
3. Updates itself based on new information.

Advantages and Disadvantages

Advantages of ML based Diagnosis systems

1. Improved Accuracy: Machine learning models can analyze vast amounts of data and provide accurate diagnoses based on patterns and trends that may be missed by human doctors.

Advantages of ML based Diagnosis systems

1. Improved Accuracy: Machine learning models can analyze vast amounts of data and provide accurate diagnoses based on patterns and trends that may be missed by human doctors.
2. Early Detection: By analyzing patient data in real-time, machine learning models can detect potential health issues at an early stage, enabling prompt and effective treatment.

Advantages of ML based Diagnosis systems

1. Improved Accuracy: Machine learning models can analyze vast amounts of data and provide accurate diagnoses based on patterns and trends that may be missed by human doctors.
2. Early Detection: By analyzing patient data in real-time, machine learning models can detect potential health issues at an early stage, enabling prompt and effective treatment.
3. Personalized Treatment: Machine learning models can analyze patient data and provide personalized treatment plans, taking into account individual factors such as age, gender, and medical history.

Advantages of ML based Diagnosis systems

1. Improved Accuracy: Machine learning models can analyze vast amounts of data and provide accurate diagnoses based on patterns and trends that may be missed by human doctors.
2. Early Detection: By analyzing patient data in real-time, machine learning models can detect potential health issues at an early stage, enabling prompt and effective treatment.
3. Personalized Treatment: Machine learning models can analyze patient data and provide personalized treatment plans, taking into account individual factors such as age, gender, and medical history.
4. Increased Efficiency: With the use of machine learning models, healthcare providers can make more efficient use of their time, reducing wait times for patients.

Disadvantages of ML based Diagnosis systems

Presence of selection bias in training data

Medical diagnoses involve highly personal information, which can result in selection bias in the data used to train our models. To address this issue, we suggest two remedies:

1. **Short-term remedy:** Use matching techniques to minimize selection bias.

Disadvantages of ML based Diagnosis systems

Presence of selection bias in training data

Medical diagnoses involve highly personal information, which can result in selection bias in the data used to train our models. To address this issue, we suggest two remedies:

1. **Short-term remedy:** Use matching techniques to minimize selection bias.
2. **Long-term remedy:** Obtain anonymous data at the time of diagnosis to eliminate initial bias.

Disadvantages of ML based Diagnosis systems

Anchoring bias

Psychological constraints such as anchoring bias can affect diagnostic accuracy. To mitigate this issue, we suggest providing a small list of highly likely diagnoses in a randomized order. This approach can speed up the diagnostic process for medical professionals while avoiding the negative effects of anchoring.

Disadvantages of ML based Diagnosis systems

Anchoring bias

Psychological constraints such as anchoring bias can affect diagnostic accuracy. To mitigate this issue, we suggest providing a small list of highly likely diagnoses in a randomized order. This approach can speed up the diagnostic process for medical professionals while avoiding the negative effects of anchoring.

One can argue that diagnosis after a machine prediction is made is equivalent to getting a second opinion: therefore there is as much anchoring as there would be in most general cases.

Naive roadmap

Naive Roadmap

1. Using a mixed mode: First use decision trees to segregate influencing factors such as demography, gender(?) and age, then using a model based on the leaf reached.

Naive Roadmap

1. Using a mixed mode: First use decision trees to segregate influencing factors such as demography, gender(?) and age, then using a model based on the leaf reached.
2. Method of proper outlier analysis needed.

Naive Roadmap

1. Using a mixed mode: First use decision trees to segregate influencing factors such as demography, gender(?) and age, then using a model based on the leaf reached.
2. Method of proper outlier analysis needed.
3. Models can be parametric in low depth and unparametric in high depth. On can also go for efficient matching techniques(similar to KNN).

Topics on which we need more Information

1. In cases where the preliminary symptoms don’t give conclusive results, how are further diagnosis steps taken?
2. Are symptoms dependent on demography/skin colour? How conclusively can demography be determined? How are mixed-racial patients treated?
3. How to determine cost function? Is it disease specific?