Hidden Markov Model for analyzing time-series health checkup data
Overview

Introduction
- background
- related works
- the purpose of our work

Methods
- HMM

Experiments
- model training
- risk estimation

Conclusion
1. Introduction
Background

Recent medical issues in Japan:
- Rapid aging of Japanese society
- Increase of medical expenses

Early detection of lifestyle-related diseases
- Hypertension, hypercholesterolemia, and diabetes mellitus
- Prevention of the progress of these diseases

Health checkup data and their analysis are necessary.
Related Works

Regression analysis, decision trees
- to discover regularity
- to predict disease development
- Most studies used one-time health checkup records

But

Temporal information is much important to analyze the progress, recovery and aggravation of the diseases.
Purpose

To make a model which can deal with the temporal progress of health checkup data

We propose to use HMM as a statistical model.

- HMM can model health condition change.
- HMM can predict future health condition.
- We investigate model parameters of trained HMMs.
- We compare state properties with manual examination of health risk examination.
2. Methods
Data Overview

Training data
- health checkup records provided by the medical center in Gifu prefecture (2002-2007)
- use the records with continuous six-year inspection data
- chose 8 parameters from the records

<table>
<thead>
<tr>
<th>8 parameters</th>
<th>age</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>51</td>
<td>52</td>
<td>53</td>
</tr>
<tr>
<td>1. Body mass index (kg/m²)</td>
<td>BMI</td>
<td>22.1</td>
<td>21.7</td>
</tr>
<tr>
<td>2. Systolic blood pressure (mmHg)</td>
<td>SBP</td>
<td>130</td>
<td>132</td>
</tr>
<tr>
<td>3. Hematocrit (%)</td>
<td>Ht</td>
<td>44.8</td>
<td>44.2</td>
</tr>
<tr>
<td>4. Platelet (10⁴/ml)</td>
<td>PLT</td>
<td>16.7</td>
<td>16.9</td>
</tr>
<tr>
<td>5. Glucose oxidase test (IU/l)</td>
<td>GOT</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>6. Total cholesterol (mg/dl)</td>
<td>T.Chol</td>
<td>266</td>
<td>263</td>
</tr>
<tr>
<td>7. Neutral fat (mg/dl)</td>
<td>TG</td>
<td>148</td>
<td>169</td>
</tr>
<tr>
<td>8. Casual blood glucose (mg/dl)</td>
<td>CBG</td>
<td>119</td>
<td>176</td>
</tr>
</tbody>
</table>
Merits of Hidden Markov Model

- HMM has **flexibility on temporal changes**.
  - Health checkup data are obtained year by year.
  - The change depends on each person.

- HMM has **transition probabilities**.
  - The transition probability can be used to estimate ...  
    - the risk of disease development,  
    - the possibility of recovery.

- HMM is structured by **many states**.
  - State parameters (mean values) are useful for the classification by the risk level.
The structure of HMM

- The HMM has six states, structured in 2x3.
- In the left and right columns, all the transitions are allowed.
- Transitions in every rows are possible.
3. Experiments
# Model Training

**experiment condition**

<table>
<thead>
<tr>
<th>Data Source</th>
<th>health checkup records provided by the medical center in Gifu prefecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Data</td>
<td>male 30’s : 4,164                                               female 40’s : 2,480</td>
</tr>
<tr>
<td></td>
<td>male 40’s : 5,733                                               female 50’s : 3,481</td>
</tr>
<tr>
<td></td>
<td>male 50’s : 7,604</td>
</tr>
<tr>
<td>Parameters</td>
<td>BMI,SBP,Ht,PLT,GOT,T.Chol,TG,CBG</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Baum-Welch algorithm (EM algorithm) flat start</td>
</tr>
</tbody>
</table>

- We evaluate HMMs by **investigating state parameters and transitions in the HMMs** using training data.
- We investigate states in HMMs by **comparing mean values in the states with the manual labels**.
HMM parameters for male 40’s

Normalized mean values in each state

BMI  SBP  GOT  TG

state1

state2

state3

state4

state5

state6

healthy

low-grade unhealthy

high-grade unhealthy
Transition probabilities of male 40’s

- Self-loop transition have larger probabilities than others.
- Transitions in the same row are followed.
- Female HMMs are similar to male HMMs.
Health risk examination

Training data have health risk examination results.

Classification of health risk examination in Japan

A : almost no risk
B : small risk
C : follow-up is necessary
D1 : recommendation of treatment
D2 : recommendation of treatment and detailed inspection

We investigated the tendency of the health risk examination in each state.
Health risk examination results

![Pie charts showing health risk examination results for males in 30's, 40's, and 50's.](image)
Discussion

The states of HMM are classified into the three groups

- **Healthy states** (state1 & state2)
  - Mean vectors have low values.
  - This group indicates ‘healthy’ or ‘normal’ conditions.
- **Low-grade unhealthy states** (state3 & state4)
  - Mean vectors are relatively high and sometimes exceed the upper limit.
  - This group indicates necessity of successive monitoring is essential.
- **High-grade unhealthy states** (state5 & state6)
  - Mean vectors are out of the acceptable range.
  - This group indicates necessity of treatment of life-style diseases.
Discussion

- HMMs can model health condition and health risk.

- Both the mean vector investigation and the health risk examination results show the same classification tendency.

  - The self-loop transitions and transitions to next state in the same row are larger.

  - Usually there are few drastic changes in only one year, and in most cases the inspection result is similar to the prior year.

  - It would be informative if a state transition indicates a change of health risk level.
4. Conclusion
Conclusion

- We proposed the introduction of HMM to health checkup data analysis.

- HMM which was built is suitable for categorizing three risk levels.

- HMM shows the possibility of estimating the risk of lifestyle-related diseases.

- HMM must be optimized for the data and the task.
Future work

- Investigation of current risk estimation of lifestyle-related diseases by using our approaches

- Estimation of future health risk by using HMMs and checkup data neighboring to the target data.

- Investigation of influence of factors for HMM parameters and health risk examination
Thank you for your kind attention