Chapter XLV
Application of Text Mining Methodologies to Health Insurance Schedules

Ah Chung Tsoi
Monash University, Australia

Phuong Kim To
Tedis P/L, Australia

Markus Hagenbuchner
University of Wollongong, Australia

ABSTRACT

This chapter describes the application of a number of text mining techniques to discover patterns in the health insurance schedule with an aim to uncover any inconsistency or ambiguity in the schedule. In particular, we will apply first a simple "bag of words" technique to study the text data, and to evaluate the hypothesis: Is there any inconsistency in the text description of the medical procedures used? It is found that the hypothesis is not valid, and hence the investigation is continued on how best to cluster the text. This work would have significance to health insurers to assist them to differentiate descriptions of the medical procedures. Secondly, it would also assist the health insurer to describe medical procedures in an unambiguous manner.

AUSTRALIAN HEALTH INSURANCE SYSTEM

In Australia, there is a universal health insurance system for her citizens and permanent residents. This publicly-funded health insurance scheme is administered by a federal government department called

the Health Insurance Commission (HIC). In addition, the Australian Department of Health and Ageing (DoHA), after consultation with the medical fraternity, publishes a manual called Medicare Benefit Schedule (MBS) in which it details each medical treatment procedure and its associated rebate to the medical service providers who provide such services. When a patient visits a medical service provider, the HIC will refund or pay the medical service provider at the rate published in the MBS[1] (the MBS is publicly available online from http://www.health.gov.au/pubs/mbs/mbs/css/index.htm).

Therefore, the description of medical treatment procedures in the MBS should be clear and unambiguous to interpretation by a reasonable medical service provider as ambiguities would lead to the wrong medical treatment procedure being used to invoice the patient or the HIC. However, the MBS has developed over the years, and is derived through extensive consultations with medical service providers over a lengthy period. Consequently, there may exist inconsistencies or ambiguities within the schedule. In this chapter, we propose to use text mining methodologies to discover if there are any ambiguities in the MBS.

The MBS is divided into seven categories, each of which describes a collection of treatments related to a particular type, such as diagnostic treatments, therapeutic treatments, oral treatments, and so on. Each category is further divided into groups. For example, in category 1, there are 15 groups, A_1, A_2, …, A_{15}. Within each group, there are a number of medical procedures which are denoted by unique item numbers. In other words, the MBS is arranged in a hierarchical tree manner, designed so that it is easy for medical service providers to find appropriate items which represent the medical procedures provided to the patient.[2] This underlying MBS structure is outlined in Figure 1.

This chapter evaluates the following:

Figure 1. An overview of the MBS structure in the year of 1999
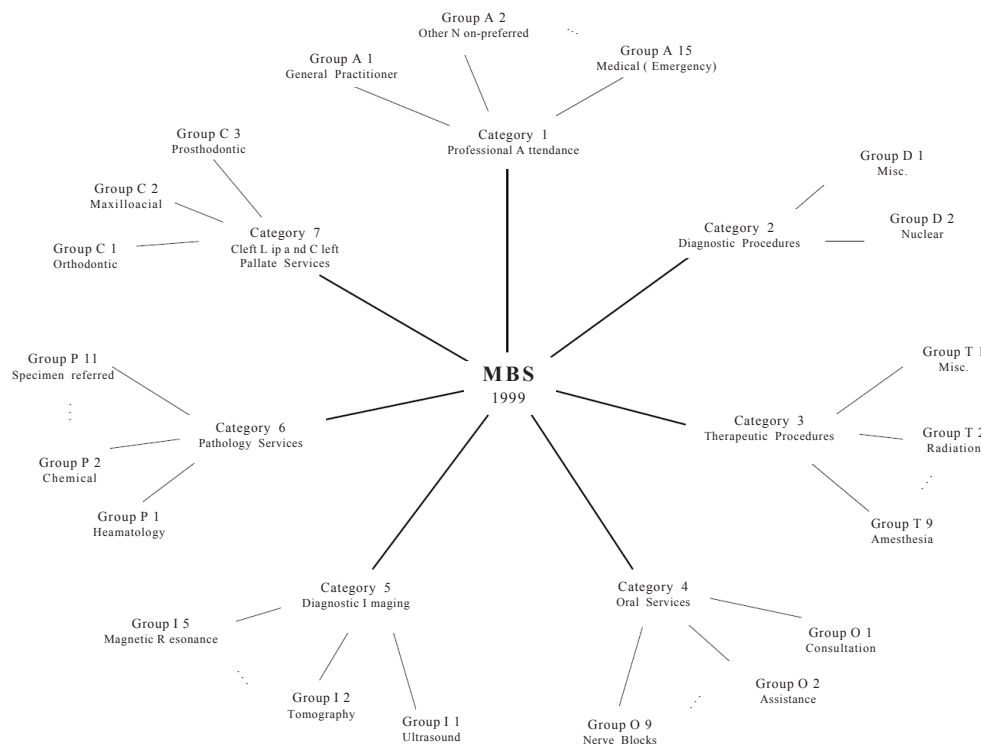
Related Content

Efficient Computation of Top-K Skyline Objects in Data Set With Uncertain Preferences

Nitesh Sukhwani, Venkateswara Rao Kagita, Vikas Kumarand Sanjaya Kumar Panda (2021). *International Journal of Data Warehousing and Mining (pp. 68-80).*

www.irma-international.org/article/efficient-computation-of-top-k-skyline-objects-in-data-set-with-uncertain-preferences/286616

Mining for Mutually Exclusive Items in Transaction Databases

George Tzanisand Christos Berberidis (2007). *International Journal of Data Warehousing and Mining (pp. 45-59).*

www.irma-international.org/article/mining-mutually-exclusive-items-transaction/1789

Application of Machine Learning Techniques for Railway Health Monitoring

G.M. Shafiullah, Adam Thompson, Peter J. Wolfsand A.B.M. Shawkat Ali (2010). *Dynamic and Advanced Data Mining for Progressing Technological Development: Innovations and Systemic Approaches (pp. 396-421).*

www.irma-international.org/chapter/application-machine-learning-techniques-railway/39650

Opinion Mining of Twitter Events using Supervised Learning

Nida Hakakand Mahira Kirmani (2022). *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines (pp. 382-397).*

www.irma-international.org/chapter/opinion-mining-of-twitter-events-using-supervised-learning/308498

A Survey on Temporal Data Warehousing

Matteo Golfarelliand Stefano Rizzi (2009). *International Journal of Data Warehousing and Mining (pp. 1-17).*

www.irma-international.org/article/survey-temporal-data-warehousing/1820