

Hospital Recommendation System

Silvanus David, 140953350, Himank Maan, 140953356, and Pratyay Amrit, 140953430

Abstract—Recommendation systems have been flooding the IT industry ever since the concept of information retrieval came into existence. Every web application uses one of these recommendation systems, which helps users of said applications to get personalized suggestions on the basis of their past records. The Hospital Recommendation System aims to use this concept for hospitals to get an idea of where to set up health camps around the area. A hospital maintains a record of every patient that visits the hospital with an issue. This record includes details such as disease, symptoms, addresses etc. **If this data can be classified on the basis of this data**, it becomes easy for the hospital to understand which locality is prone to which disease and will allow these hospitals to set up health camps with focus on these diseases, thus eliminating diseases before they become an issue. **With the help of doctors, this data can be used to inform the concerned authorities of the reasons why the area is prone to the particular disease.** For example, a leaked drainage can invite disease causing insects in the area. If this data is provided and wisely used, health problems can be greatly reduced. This paper aims to find ways to obtain this data.

Keywords—Recommendation Systems, Information Retrieval, Medical Recommendation, Hospital Recommendation

I. INTRODUCTION

The outburst of population in recent years have brought all institutions under great pressure. Hospitals, in particular, are flooded with patients everyday. Even then, to the rural citizens, accessibility is, and has always been a major issue. To counter this, hospitals have been periodically setting up health camps for free health checkups and to raise general awareness. It is often noticed that, these health camps diagnose similar kind of issues in their respective regions. Using data mining and classification technologies, we can identify these medical hotspots prior to visiting them. Efficient extraction of valuable information from hospitals can be put to greater use, attacking diseases from many more directions. This can be achieved in many ways, depending on the type of data available in the hospitals. It is interesting to note that some of these goals require nothing more than a better visualization of data. We have come up with the following cases where the hospital would benefit from our efforts:

1. To classify places as prone to diseases and set up health camps.
2. To identify which disease is more common in these prone locations to provide facilities more focussed to the cause.
3. To identify the cause of these common diseases and report the same to concerned authorities.
4. To classify a new patient coming from an identified area as more susceptible to the common disease and have him/her diagnosed for the same.

These cases are but a few ways this system can help the hospitals by providing a recommendation system.

A. Data Set

For the purpose of demonstration of the usability of these systems, we will use a data set acquired from Data World^[1]. This dataset is for death of California residents, by cause of death, by Zip Code, for the years 1999-2013. The death counts per cause of death, represents the deaths to California residents. A non-California resident who dies inside of California is not included in the dataset.

year	zip	cod	count	x	y
1999	90009	LIV	1	33.945	-118.383
1999	90068	DIA	1	34.116	-118.330
1999	90704	INJ	1	33.332	-118.344
1999	90813	ALZ	1	33.782	-118.183
1999	91024	INJ	1	34.165	-118.052
1999	91202	HOM	1	34.165	-118.266
1999	91301	PNF	1	34.157	-118.757
1999	91354	ALZ	1	34.447	-118.537
1999	91377	PNF	1	34.136	-118.774
1999	91765	ALZ	1	34.007	-117.810
1999	91905	LIV	1	32.672	-116.320
1999	92041	PNF	1	?	?
1999	92242	ALZ	1	34.149	-114.339
1999	92251	DIA	1	32.847	-115.573
1999	92257	LIV	1	33.378	-115.696
1999	92259	STK	1	32.739	-115.993
1999	92270	HOM	1	33.764	-116.422

Fig. 1. Sample data set. The 'x' and 'y' attributes signify the geographical coordinates of the location of the zip code. The data set is a total of 10000 entries

B. Methodology

For each of the cases mentioned earlier, we will be suggesting ways to achieve the goal, and we will apply these ways on the data set. Some of them just involve a better visualization of the dataset. We will be comparing the different ways used in each of these methods. For example, for the first case, to identify the places prone to diseases in general would require a simple place vs disease graph. However, clustering the data would provide a better idea of the geographical boundaries

prone to the diseases. Pros and cons will be discussed in each of the cases in their respective sections. To generate results we will be using RapidMiner Studio for the purpose of demonstration. If a hospital recommendation system is to be developed, it can be developed using any of the programming technologies that the developer feels comfortable in.

II. CONCLUSION

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

APPENDIX A

PROOF OF THE FIRST ZONKLAR EQUATION

Some text for the appendix.

ACKNOWLEDGMENT

The authors would like to thank...

REFERENCES

- [1] H. Kopka and P. W. Daly, *A Guide to L^AT_EX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.