MARMARA UNIVERSITY

FACULTY OF ENGINEERING

2020-2021 SPRING SEMESTR

CSE4062

Introduction to Data Science and Analytics

Group 6

Proposal

Investigation of Post-Chernobyl Papillary Thyroid Cancer Gene Expression Levels

150816004 Defne ÇIĞ, Bioengineering - defnecig@marun.edu.tr

150816019 Kasım Anlatır Bioengineering - kasimanlatirr@gmail.com

150216031 Cem Telliağaoğlu Environmental Engineering - cem.telliagaoglu@gmail.com

199520021 Lukas Eckerle Industrial Engineering - lukas.eckerle@gmail.com

150116061 Ertuğrul Sağdıç Computer Engineering - ertugrulsagdic98@gmail.com

Description of the Project

The thyroid gland is an organ located under the thyroid cartilage in the front of the neck and has an endocrine function. A type of cancer caused by the transformation of cells in the thyroid gland into cancer is called thyroid cancer and it has 4 types: papillary, follicular, medullary, and anaplastic. Of these, the most commonly seen is papillary thyroid cancer (PTC) [1].

In addition to many known causes of thyroid cancer, thyroid glands are one of the most affected areas in our body by radiation. After the nuclear catastrophe in Chernobyl, the increase in cancer rates in the surrounding areas is remarkable. Understanding the cancer mechanism seen in people living in the surrounding areas is important for future studies [2].

The aim of this project is to detect changes in the genome level that may be caused by Chernobyl radiation. For this reason, samples taken from patients known to have been exposed to radiation will be compared with samples taken from patients known to have not been exposed to radiation. The data to be studied includes (GSE35570) 33 cases exposed to Chernobyl radiation, 32 cases not exposed to Chernobyl radiation and 51 healthy cases [3]. Primarily, differentially expressed genes (DEGs) will be detected for both groups. Then, the presence of genes that can distinguish these two groups will be investigated. At the end of the study, it is aimed to determine whether a thyroid cancer is caused by radiation or not.

Data Features

Our dataset includes three groups. The first of these is the ones exposed to the radio from Chernobyl and contains 54676 rows and 34 columns. Another group includes those known not to be exposed to Chernobyl radiation and contains 54676 rows and 33 columns. The last group represents the healthy control group and contains 54676 rows and 52 columns. The ID_Ref column shows the ID assigned to each gene, and each column represents one patient. Our data shows us the expression values of each gene of each patient.

Our goal is to find out if there are differences between the gene expressions of the cancer that was exposed to radiation and the genes of the cancer that was not exposed to radiation, and if so, which of these genes are different. To reach this goal we compare the genes of the exposed cancer with the control group without cancer and do the same with the genes of the non-exposed exposed cancer by calculating the p-values in order to either reject or accept the Null hypothesis. Then we compare the genes of the radiation-exposed cancer with the genes of the non-exposed cancer for which we could accept the Null hypothesis to find the genes which change for radiation-induced cancer.
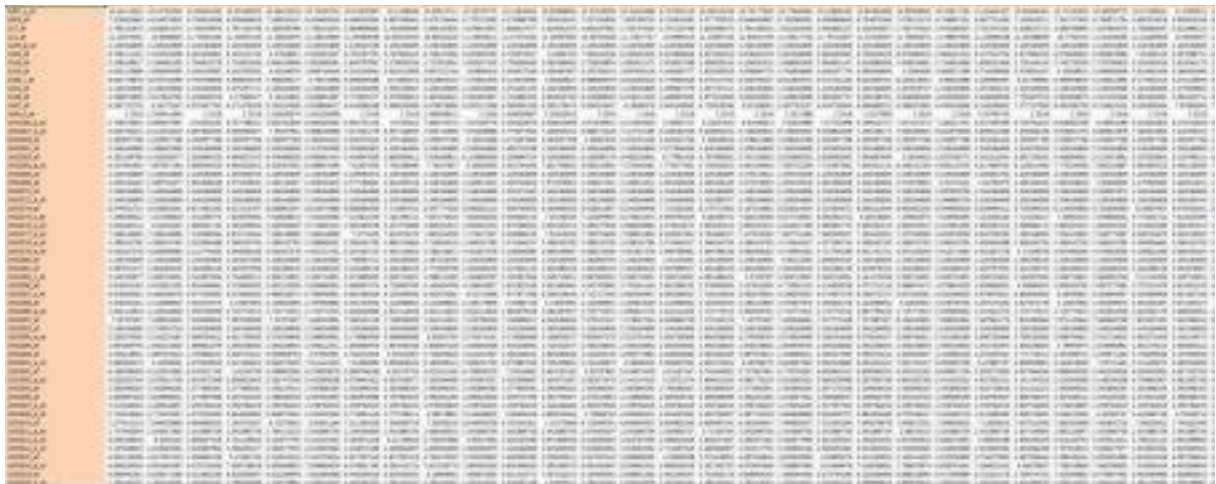
Figure 1: Gene expressions of the exposed to Chernobyl radiation
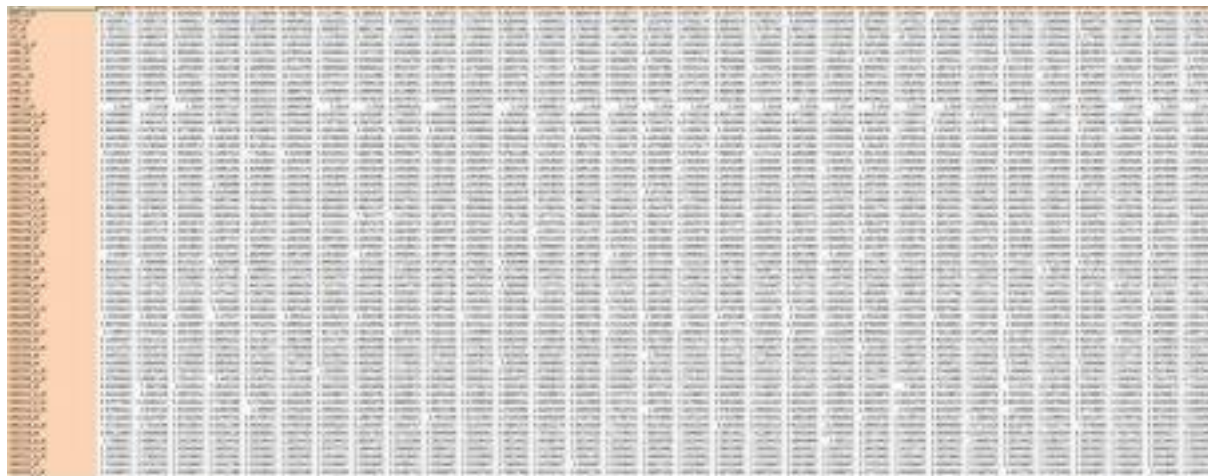
Figure 2: Gene expressions of the non-exposed to Chernobyl radiation

REFERENCES

1. Cabanillas, M. E., McFadden, D. G., & Durante, C. (2016). Thyroid cancer. *The Lancet*, *388*(10061), 2783-2795.

2. Ron, E., Lubin, J. H., Shore, R. E., Mabuchi, K., Modan, B., Pottern, L. M., ... & Boice Jr, J. D. (1995). Thyroid cancer after exposure to external radiation: a pooled analysis of seven studies. *Radiation research*, *141*(3), 259-277.

3. Handkiewicz-Junak, D., Swierniak, M., Rusinek, D., Oczko-Wojciechowska, M., Dom, G., Maenhaut, C., ... & Jarzab, B. (2016). Gene signature of the post-Chernobyl papillary thyroid cancer. *European journal of nuclear medicine and molecular imaging*, *43*(7), 1267-1277.