**Lung Cancer Stage prediction with machine learning**

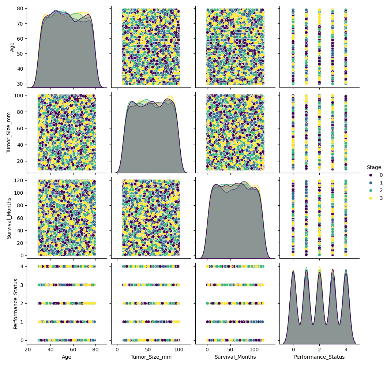
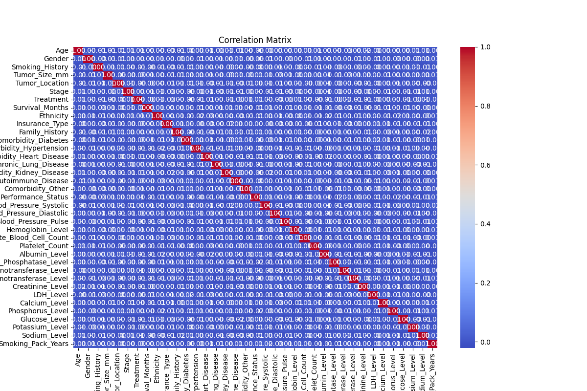
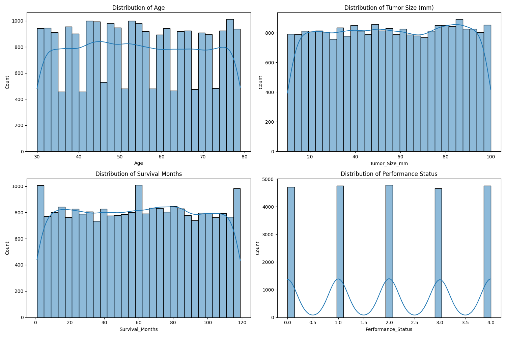
**Vladimir Gološin, Kristina Andrijin, Stefan Bogdanović**

1. Motivation

Doctors are in short supply and those who are available are often overworked and tired, especially the lung cancer specialists. In order to help them out a machine learning model can be created to aid in cancer diagnostics.

2. Research questions

Specifically, we want to make a machine learning model for classification of lung cancer stages based on various parameters, these include: Age, Gender, Smoking\_History, Tumor\_Size\_mm, Tumor\_Location, Stage, Treatment, Survival\_Months, Ethnicity, Insurance\_Type, Family\_History, Comorbidity\_Diabetes, Comorbidity\_Hypertension, Comorbidity\_Heart\_Disease, Comorbidity\_Chronic\_Lung\_Disease, Comorbidity\_Kidney\_Disease, Comorbidity\_Autoimmune\_Disease, Comorbidity\_Other, Performance\_Status, Blood\_Pressure\_Systolic, Blood\_Pressure\_Diastolic, Blood\_Pressure\_Pulse, Hemoglobin\_Level, White\_Blood\_Cell\_Count, Platelet\_Count, Albumin\_Level, Alkaline\_Phosphatase\_Level, Alanine\_Aminotransferase\_Level, Aspartate\_Aminotransferase\_Level, Creatinine\_Level, LDH\_Level, Calcium\_Level, Phosphorus\_Level, Glucose\_Level, Potassium\_Level, Sodium\_Level, Smoking\_Pack\_Years. We will now take a look at how these are distributed among the cancer stages.



We can see that the data is very equally distributed among the classes. Besides that, the data has a very low correlation with each other, it is around 0. In the final set of graphs we can see that the data has a rather uniform distribution, which means it is most likely synthetic in origin.

3. Related work

<Discuss how the others addressed that problem.>

4. Methodology

<Describe HOW you did address the problem. >

5. Discussion

<Describe the experiment (testing procedure, evaluation measure and hyperparameter optimization). Describe obtained results, including error analysis.>

6. References

<List references.>

<The maximum number of pages is 4>