**Title: Development of a system in order to predict tumor metastasis and recurrence and to improve diagnostics and treatment of cancer patients**

**Problem**

One of the most popular disease in our modern society is cancer. Many people die because of the consequences of tumors, evolving in their bodies and capturing nearly all of their good, healthy cells. The cause of tumors is the mutation of cells into 'bad' cells which clone themselves uncontrolled and irregularly. This makes it very hard to diagnose and prevent new tumors. Sometimes, even experts cannot diagnose precisely, e.g. in case of malignant tumors which are not delimited in the MRI pictures but tend to proliferate in an unknown degree. Another problem is the point of time the diagnosis is made. Since many people who do not suffer from any chronic disease tend to go to the doctor infrequently, the tumor is be detected too late. What is more, most biopsies take up to multiple days which is a long time.

**Goal**

Given the above explained problems, the goals are to improve diagnostic and treatment of tumor patients by developing a model which predicts tumor metastasis and recurrence. Moreover, the system shall help doctors to quickly identify a patient whose cells are mutating so that he can intervene and treat the patient at the right time. Furthermore, the information process shall be improved by setting up a mobile and web application which keeps the patients up to date and serves as a communication base for healthcare providers.

**Methodology**

First, to get an overall business understanding, the key users (patients and treating doctors) are interviewed to find out the scope of the mobile and web application. At this stage, design thinking methods are used. Second, CRISP-DM (CRoss-Industry Standard Process for Data Mining) is used for data analysis process. Beside business understanding, it includes data understanding, data preparation, modelling, evaluation and deployment. Third, pattern recognition is implemented in order to find out patterns in the mutations of tumor cells. The pattern recognition includes feature extraction (e.g. tumor cell vs. healthy cell values) and classification. When characterizing attributes common to all patterns belonging to a class, genetic algorithms are used. Lastly, to develop the mobile application, the framework IONIC is implemented since it scales both web and mobile applications.

**Benefits/Contribution**

By reason of the adopted methodologies, there are many benefits for all users. To be more precise, patients do not have to suffer from the treatment of cytostatics too long since the algorithm predicts the exact location of the tumor and only a local treatment is needed. Besides, patients do not have to wait too long for the treatment outcome since they are informed by the mobile application.

**Relevance**

To conclude, this master thesis is very relevant in both scientific and economic contexts. Regarding the scientific context, diagnostic and treatment of cancer patients are improved. Human errors, such as unprecise detection of tumors in MRI images are eliminated. Moreover, patients feel more comfortable since they can stay longer at home.

Relating to the economic context, the developed system reduces the costs of chemotherapies and treatments since they can be personalized for each patient. After that, doctors are unburdened and can treat more patients at the same time. Besides, there are more free beds in a hospital.

Vom Prof vorgeschlagene Gliederung:

* **Forschungsfrage/Zielsetzung der Arbeit**
* **Methodisches & inhaltliches Vorgehen**
* **Kurzüberblick über problemrelevante Literatur**
* **Erste Gliederung der Arbeit**

1. **Use and scope of XXX**

**1.1. Motivation**

**1.2. Aim and Scope**

1. **Related Work**
   1. **Current Algorithms/Solutions to predict metastasis**
   2. **Conflicts**
2. **Functionality of XXX, YYY**

**3.1. Technology XXX**

**3.2. Technology YYY**

1. **User Research**

**4.1. User interviews**

**4.2. What is important? – Relevant features**

**4.3. How can processes be improved?**

1. **Development of system to predict metastasis and recidives**

**5.1. Data Research**

**5.2. CRISP-DM: Model Planning and Learning**

**5.3. Testing**

**5.3. Results**

**6. Results**

**6.1. Validation of results**

**6.2. Limitations in the**

**7. Conclusion and Outlook**

ResearchGate. „(2) (PDF) Volumetric Mapping of Intra‐ and Extracellular PH in the Human Brain Using 31P MRSI at 7T“. Zugegriffen 6. April 2020. https://www.researchgate.net/publication/340364396\_Volumetric\_mapping\_of\_intra-\_and\_extracellular\_pH\_in\_the\_human\_brain\_using\_31P\_MRSI\_at\_7T.

IBM Research Blog. „AI Models Predict Breast Cancer with Radiologist-Level Accuracy“, 18. Juni 2019. https://www.ibm.com/blogs/research/2019/06/ai-models-radiologist-level-accuracy/.

Batmanov, Kirill, Jan Delabie, und Junbai Wang. „BayesPI-BAR2: A New Python Package for Predicting Functional Non-Coding Mutations in Cancer Patient Cohorts“. *Frontiers in Genetics* 10 (2019). https://doi.org/10.3389/fgene.2019.00282.

Benzekry, Sébastien, Clare Lamont, Afshin Beheshti, Amanda Tracz, John M. L. Ebos, Lynn Hlatky, und Philip Hahnfeldt. „Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth“. *PLOS Computational Biology* 10, Nr. 8 (28. August 2014): e1003800. https://doi.org/10.1371/journal.pcbi.1003800.

BMBF-Internetredaktion. „Karliczek: Künstliche Intelligenz für Krebsbehandlung nutzen - BMBF“. Bundesministerium für Bildung und Forschung - BMBF. Zugegriffen 26. März 2020. https://www.bmbf.de/de/karliczek-kuenstliche-intelligenz-fuer-krebsbehandlung-nutzen-11047.html.

„Cancer Facts & Figures 2016 | American Cancer Society“. Zugegriffen 17. April 2020. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2016.html.

Data.gov. „Data.Gov“. Zugegriffen 19. April 2020. https://www.data.gov/.

„Digitale Diagnostik für gezielte Krebstherapie“. Zugegriffen 26. März 2020. https://www.healthcare-computing.de/digitale-diagnostik-fuer-gezielte-krebstherapie-a-912839/.

„Double Reading in Breast Cancer Screening: Cohort Evaluation in the CO-OPS Trial | Radiology“. Zugegriffen 19. April 2020. https://pubs.rsna.org/doi/full/10.1148/radiol.2018171010.

Gad, Ahmed. „Genetic Algorithm Implementation in Python“. Medium, 13. Juni 2019. https://towardsdatascience.com/genetic-algorithm-implementation-in-python-5ab67bb124a6.

„GHO“. Zugegriffen 19. April 2020. https://www.who.int/data/maternal-newborn-child-adolescent/monitor.

Herman, Alexander B., Van M. Savage, und Geoffrey B. West. „A Quantitative Theory of Solid Tumor Growth, Metabolic Rate and Vascularization“. *PLoS ONE* 6, Nr. 9 (29. September 2011). https://doi.org/10.1371/journal.pone.0022973.

„Human Mortality Database“. Zugegriffen 19. April 2020. https://www.mortality.org/.

„Introduction to Optimization with Genetic Algorithm“. Zugegriffen 20. März 2020. https://www.linkedin.com/pulse/introduction-optimization-genetic-algorithm-ahmed-gad.

Mallawaarachchi, Vijini. „Introduction to Genetic Algorithms — Including Example Code“. Medium, 1. März 2020. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3.

Mayo, Charles S., Marc L. Kessler, Avraham Eisbruch, Grant Weyburne, Mary Feng, James A. Hayman, Shruti Jolly, u. a. „The Big Data Effort in Radiation Oncology: Data Mining or Data Farming?“ *Advances in Radiation Oncology* 1, Nr. 4 (1. Oktober 2016): 260–71. https://doi.org/10.1016/j.adro.2016.10.001.

ResearchGate. „(PDF) Applications of Machine Learning in Cancer Prediction and Prognosis“. Zugegriffen 19. April 2020. http://dx.doi.org/10.1177/117693510600200030.

ResearchGate. „(PDF) Applying Genetic Algorithms on Pattern Recognition: An Analysis and Survey“. Zugegriffen 5. März 2020. http://dx.doi.org/10.1109/ICPR.1996.546921.

„Search | HealthData.gov“. Zugegriffen 19. April 2020. https://healthdata.gov/search/type/dataset?query=cancer.

Sun, Yingshuai, Sitao Zhu, Kailong Ma, Weiqing Liu, Yao Yue, Gang Hu, Huifang Lu, und Wenbin Chen. „Identification of 12 Cancer Types through Genome Deep Learning“. *Scientific Reports* 9, Nr. 1 (21. November 2019): 1–9. https://doi.org/10.1038/s41598-019-53989-3.

tien-tang. *tien-tang/tumor-classification\_growth-rate*. MATLAB, 2020. https://github.com/tien-tang/tumor-classification\_growth-rate.