FOM - Hochschule für Oekonomie & Management
Hamburg

Master-Studiengang Big Data & Business Analytics
5. Semester

Development of a solution to detect skin lymphoma and to improve treatment processes

Betreuer: Prof. Dr. David Matusievicz

Autor: Jacqueline Franßen

Matrikel-Nr: 496804

5. Fachsemester

Hamburg, den 28.02.2021

Contents

List of Figures

List of Tables

1 Abstract

This scientific article focusses on the development of a neural network to detect tumors and to prevent skin cancer from recurring. The main purpose of this scientific article is to develop... This can lead to... What is more, the mobile app should connect both patients and healthcare practitioners by digitalizing all measured data and saving them centrally on a secured database. The first business case is that patients are frequently reminded of going to a control appointment. Second, by using their smartphone's hardware, especially the camera can be used to take a photo of potential skin cancers. Furthermore, ... Third,... Fourth,... Last, ...

2 Introduction

2.1 Problem statement

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 Magnetic Resonance Imaging (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. According to.... Another problem are All of these use cases can be automated and implemented by an application....

2.2 Aim and scope of this work

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.

The first aim of this scientific work is to develop a solution ... What is important, the

developed model is only a reference model....

3 Fundamentals

3.1 Related Work

3.1.1 IARC

The IARC is a association with the aim to raise the development within cancer treatment and research. It was formed by the World Health Organization (WHO) and has its main location in France. Besides, it provides a large library of cancerogens for users.

3.1.2 HAM10000

Dermatoscopy TBD [1]

3.1.3 CanScreen5

The IARC provides many projects based on different data sources. One large project is the CanScreen5[2] which collects all information on several cancer screening programmes across the globe. Thereby, the main goals are quality improvement and effective programme evaluation of cancer therapies. There are three types of cancer (breast, cervical and colorectal cancer) which are examined precisely. Besides, an overall graphic explains the distribution of cancer types across the worldwide population. Furthermore, the CanScreen5 project provides analysis tools to be able to see the distribution of for example breast cancer in european countries.

---

[1]Tschandl, Rosendahl, and Kittler 2018

[2]*canscreen5* 2020

3.2 Benchmarking

3.2.1 Yuma App

a mobile app to protect against sunburning

3.3 Different types of skin tumors

'The skin is made of a variety of cells, many of which are in constant motion. Round basal cells below the surface flatten as they rise to replace dead, flaking squamous cells on the surface. Melanocytes tan the skin in the sunlight, and Merkel cells give skin its ability to sense touch. When these cells become damaged, they may develop into skin cancer.' [3]

Tumor Name	Description	Benign/Malignant
Basal Cell Carcinoma	tbd	tbd
Recurrent Basal Cell Carcinoma	tbd	tbd
Squamous Cell Carcinoma	tbd	tbd
Melanoma	tbd	tbd
Karposi sarcoma (Rare Skin cancer)	tbd	tbd
Actinic keradosis (Rare Skin cancer)	tbd	tbd
Lymphoma of the skin (Rare Skin cancer)	tbd	tbd
Keratoacanthoma (Rare Skin cancer)	tbd	tbd
Merkel Cell Carcinoma	tbd	tbd

As described by Norddeutscher Rundfunk (NDR) [4], malignant skin lymphoma are caused by our immunesystem cells (also knows as lymphocytes). In such case, the lymphocytes are collected beneath the skin, grow uncontrolled and cause lymphomes. The reason for the mutation and why these cells cause skin lymphomes is unknown.

---

[3] *Types of Skin Cancer* 2018
[4] NDR 2020

3.3.1 Current diagnosis and therapy methods

When assuming skin lymphoma, doctors often examine the patient's blood or use X-ray and ultrasound to detect all tumor cells. After recognizing all malignant cells, there are several therapy methods, such as cortisone ointment, phototherapy, extracorporeal photopheresis, X-rays, antibody therapy, chemotherapy and symptomatic therapy.

Given these numerous therapy methods, it is very important to diagnose the 'right' disease. Often, lymphoma tend to mutate and hide themselves as psoriasis and are treated like them for many years. But suddenly, the doctor notices that it is not a usual psoriasis and the lymphoma can immediately lead to the patient's death. For that reason, the diagnosis process should be repeated like a cycle and should include all different data sources (e.g. not only image recognition but also blood analysis) [5].

3.3.2 Stages of skin cancer

When diagnosing tumors, a often used scale is the Tumor Node Metastasis (TNM) grading which describes the current stage of the tumor. Tumor (t) describes the tumor's size, location and how deep it has grown into the skin. Node (n) indicates whether or not the cancer cells have spread to nearby lymph nodes or the channels connecting the lymph nodes. Metastasis (m) refers to whether the cancer cells have spread to distant organs.

Stage	Tumor devel
0	'Carcinoma in situ', tumor is only present in the epidermis (uppe
1	cancer is less than 2 centimeters, about 4-5 of an inch across, has not spread
2	cancer is larger than 2 centimeters across, and has not spread to nearby organ
3	cancer has spread into facial bones or 1 nea
4	cancer can be any size and has spread (metastasized) to 1 or more lymph nodes whic

According the the detected stage, the therapy is personalized. For instance, if tumors are in stage 1, a more simple treatment is administered whereas tumors in stage 4 have to be treated using more aggressive therapies, such as chemotherapies.

---

[5]NDR 2020

Besides, there are additional criteria which can determine a higher risk for recurrence or spreading. These are for example a thickness greater than 2mm, invasion into the lower dermis or tiny nerves in the skin or a location on the ear or on a hair-bearing lip[6].

3.4 Current algorithms, solutions to recognize skin tumors and to predict metastasis

3.5 Overview: Apps to predict metastasis

There already exist some applications to predict certain types of cancer worldwide, such as [7]. These solutions provide several information about the tumors by showing them within many different types of charts. They do not specify on a certain patient (which would also cause problems with data privacy and protection) but explain them in general.

[8]

---

[6] *Types of Skin Cancer* 2018

[7] *Cancer tomorrow* 2020

[8] Mallawaarachchi 2020

4 Image Recognition and App Development

Here all used algorithms and patterns shall be explained

4.1 Image Recognition

4.2 Hybride App Development

5 Analysis and Development

5.1 Design Thinking Methods

In the following sections there are explained various kinds of design strategies to find out the real needs for the application being developed. These are adopted to two basic frameworks, the IBM Enterprise Design Thinking framework [1] as well as the Design Kit, proposed by IDEO.ORG [2]. Some of the strategies are modified a little bit and were implemented by a single person which can affect the objectivity and diversity of the methods. Nevertheless, they generated a highly usable product and many ideas for future development cycles.

5.1.1 Vertical latter

In the given figure, a vertical letter shows the initially planned goals and features of the app. The X-Axis shows the time steps whereas the Y-Axis explains the complexity of tasks. The higher on task is mentioned, the more complex it is to realize.
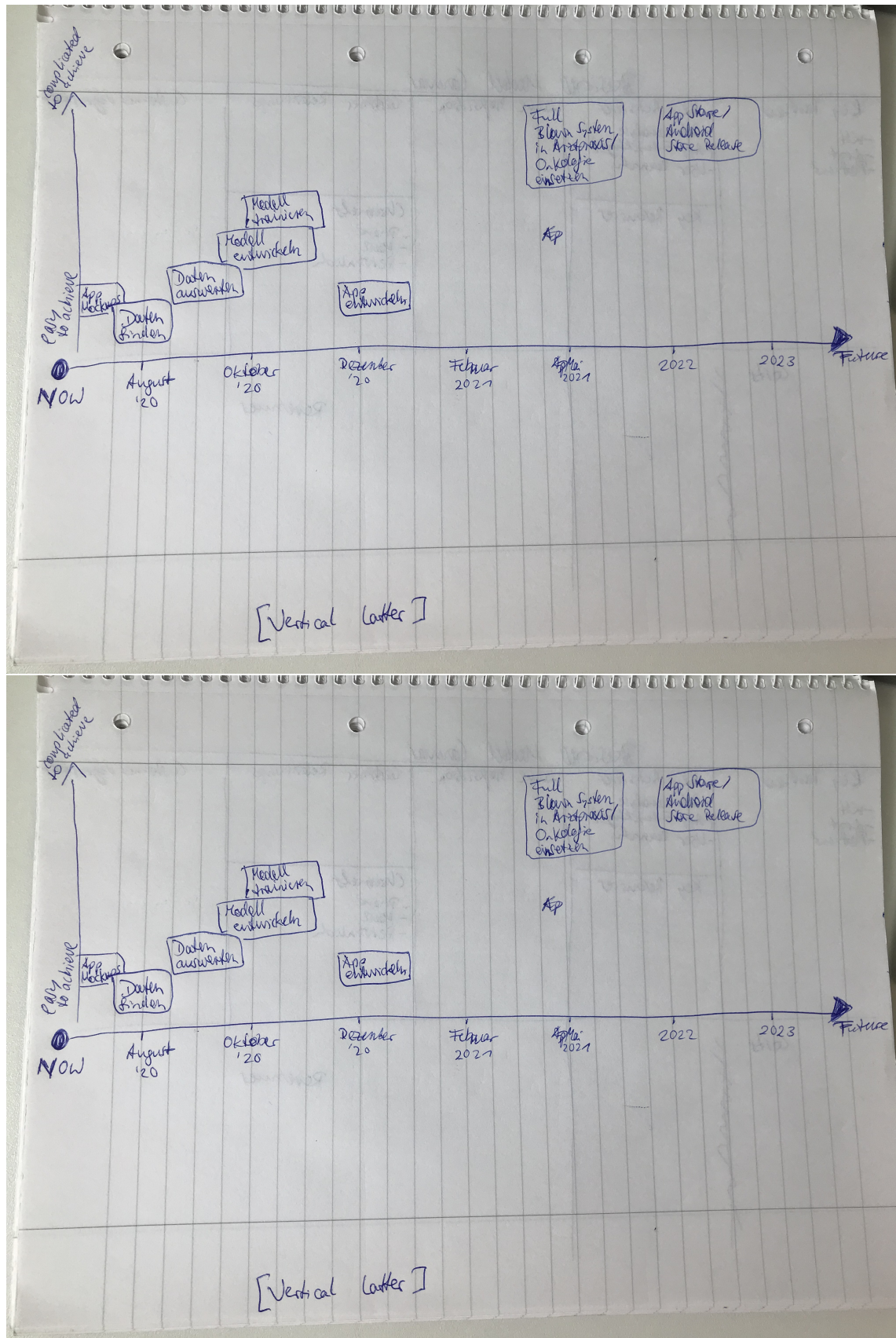
Figure 5.1: Vertical latter which defines the development goals over time

5.1.2 How might we?

5.1.3 Stakeholder Map

5.1.4 Hills

5.1.5 User Journey

5.2 User Research

5.2.1 User interviews

5.2.2 What is important? - Relevant features

5.2.3 How can processes be improved?

5.3 Development of system to predict metastasis and recidives

5.4 Data Research

5.4.1 Kaggle dataset

For training and developing the neural network, a basic Kaggle dataset was used [3]. It contains 3600 pictures with the format 224x244 px. These 3600 files are divided into benign and malignant skin cancers. For training, the neural network has only to learn

---

[1] *Design thinking activities and tools - Enterprise Design Thinking* 2018

[2] *Design Kit* 2020

[3] *Skin Cancer* 2020

how to distinguish good from bad tumors and then set a warning. The dataset did not contain a basic test set which is not bad because for testing, the data could be simply annonymized by removing the labels.

5.5 CRISP-DM: Model Planning and Learning

5.6 Testing

5.7 Results

6 Results

6.1 Validation of results

6.2 Limitation in the development process

7 Conclusion and Outlook

7.1 Conclusion

7.2 Outlook

8 Abbreviations

CRISP-DM CRoss-Industry Standard Process for Data Mining

WHO World Health Organization

IARC International Agency for Research on Cancer

HAM10000 Human Against Machine with 10000 training images

MRI Magnetic Resonance Imaging

NDR Norddeutscher Rundfunk

TNM Tumor Node Metastasis

CanScreen5 Cancer Screening in Five Continents

Bibliography

Cancer tomorrow (2020). en. Library Catalog: gco.iarc.fr. URL: `http://gco.iarc.fr/tomorrow/home` (visited on 05/24/2020).

canscreen5 (2020). Library Catalog: canscreen5.iarc.fr. URL: `https://canscreen5.iarc.fr` (visited on 05/24/2020).

Design Kit (2020). URL: `https://www.designkit.org/` (visited on 05/31/2020).

Design thinking activities and tools - Enterprise Design Thinking (May 2018). en. Library Catalog: www.ibm.com. URL: `https://www.ibm.com/design/thinking/page/toolkit` (visited on 05/31/2020).

Mallawaarachchi, Vijini (Mar. 2020). *Introduction to Genetic Algorithms — Including Example Code*. en. Library Catalog: towardsdatascience.com. URL: `https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3` (visited on 03/05/2020).

NDR (2020). *Lymphom: Wenn Tumore die Haut zerstören*. de. Library Catalog: www.ndr.de. URL: `/ratgeber/gesundheit/Lymphom-Wenn-Tumore-die-Haut-zerstoeren,lymphom102.html` (visited on 05/31/2020).

Skin Cancer (2020). *Skin Cancer: Malignant vs. Benign*. en. Library Catalog: www.kaggle.com. URL: `https://kaggle.com/fanconic/skin-cancer-malignant-vs-benign` (visited on 05/31/2020).

Tschandl, Philipp, Cliff Rosendahl, and Harald Kittler (Aug. 2018). "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions". In: *Scientific Data* 5. ISSN: 2052-4463. DOI: `10.1038/sdata.2018.161`. URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6091241/` (visited on 05/31/2020).

Types of Skin Cancer (Aug. 2018). *Types of Skin Cancer: Common, Rare and More Varieties*. en. Library Catalog: www.cancercenter.com. URL: `https://www.cancercenter.com/cancer-types/skin-cancer/types` (visited on 06/01/2020).

9 Appendix A

Ehrenwörtliche Erklärung

Hiermit versichere ich, dass die vorliegende Arbeit von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist, insbesondere dass ich alle Stellen, die wörtlich oder annähernd wörtlich aus Veröffentlichungen entnommen sind, durch Zitate als solche gekennzeichnet habe. Ich versichere auch, dass die von mir eingereichte schriftliche Version mit der digitalen Version übereinstimmt. Weiterhin erkläre ich, dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde / Prüfungsstelle vorgelegen hat. Ich erkläre mich damit nicht einverstanden, dass die Arbeit der Öffentlichkeit zugänglich gemacht wird. Ich erkläre mich damit einverstanden, dass die Digitalversion dieser Arbeit zwecks Plagiatsprüfung auf die Server externer Anbieter hochgeladen werden darf. Die Plagiatsprüfung stellt keine Zurverfügungstellung für die Öffentlichkeit dar.

Ort, Datum (Vorname Nachname)

10 Appendix B