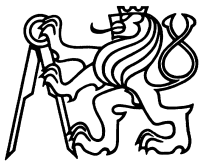
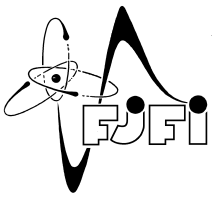


Czech Technical University in Prague
Faculty of Nuclear Sciences and Physical Engineering



Estimating patient's life expectancy after a successful kidney transplant using machine learning methods

Odhad délky života pacienta po úspěšné transplantaci ledviny pomocí metod strojového učení

Bachelor's Degree Project

Author: **Kyrylo Stadniuk**

Supervisor: **Ing. Tomáš Kouřim**

Consultant: **Ing. Pavel Strachota, Ph.D.**

Language advisor: **PaedDr. Eliška Rafajová**

Academic year: 2022/2023

Katedra: matematiky

Akademický rok: 2022/2023

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student:	Kyrylo Stadniuk
Studijní program:	Aplikovaná informatika
Název práce (česky):	Odhad délky života pacienta po úspěšné transplantaci ledviny pomocí metod strojového učení
Název práce (anglicky):	Estimating patient's life expectancy after a successful kidney transplant using machine learning methods

Pokyny pro vypracování:

1) Prozkoumejte současný přístup k transplantacím ledvin, jeho problémy a výzvy. /
Investigate the current approach to kidney transplantation, its problems and challenges.
2) Prozkoumejte příslušné metody strojového učení a metody pro hodnocení přesnosti modelu. /
Explore applicable machine learning methods and model accuracy evaluation methods.
3) Vyčistěte, předzpracujte a rozšiřte stávající datovou sadu. /
Clean, preprocess and extend the existing dataset.
4) Vytvořte prediktivní model strojového učení pro odhad délky života pacienta a ohodnoťte jeho přesnost. /
Create a predictive machine learning model estimating a patient's life expectancy and evaluate its accuracy.
5) Navrhněte úpravy skórovacího algoritmu pro transplantace ledvin na základě výsledků prediktivního modelu. /
Design an updated kidney matching compatibility scoring algorithm based on the prediction model.
6) Prozkoumejte možnost integrace dosažených výsledků do nástroje pro správu transplantací TX Matching. /
Evaluate the possibility of integrating achieved results into kidney transplantation management tool TX Matching.

Doporučená literatura:

1) P. Bruce, A. Bruce, P. Gedeck, Practical Statistics for Data Scientists, O´Reilly, 2020.
2) I. H. Witten, E. Frank, M. A. Hall, Ch. J. Pal, Data Mining : Practical Machine Learning Tools and Techniques. Morgan Kaufman, 2017.
3) A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, 2019.
4) J. J. Kim, S. V. Fuggle, S. D. Marks, Does HLA matching matter in the modern era of renal transplantation? Pediatr Nephrol 36, 2021, 31–40.
5) R. Reindl-Schwaighofer, A. Heinzel, A. Kainz, et al., Contribution of non-HLA incompatibility between donor and recipient to kidney allograft survival: genome-wide analysis in a prospective cohort. The Lancet 393, 10174, 2019, 910-917.
6) M. Wohlfahrtová, O. Viklický, R. Lischke a kolektiv, Transplantace orgánů v klinické praxi. Grada, 2021.

Jméno a pracoviště vedoucího bakalářské práce:

Ing. Tomáš Kouřim
Mild Blue, s.r.o., Plzeňská 27, Praha 5

Jméno a pracoviště konzultanta:

Ing. Pavel Strachota, Ph.D.
Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 120 00 Praha 2

Datum zadání bakalářské práce: 31.10.2022

Datum odevzdání bakalářské práce: 2.8.2023

Doba platnosti zadání je dva roky od data zadání.

V Praze dne 31.10.2022

garant oboru

vedoucí katedry

děkan

Acknowledgment:
I am grateful to Ing. Tomáš Kouřim for his expert guidance and to Dr. Pavel Strachota for his invaluable support and insightful feedback throughout this project. I would also like to extend my sincerest appreciation to PaedDr Eliška Rafajová for her language assistance.

Author's declaration:
I declare that this Bachelor's Degree Project is entirely my own work and I have listed all the used sources in the bibliography.

Prague, August 2, 2023 Kyrylo Stadniuk

Název práce:

Odhad délky života pacienta po úspěšné transplantaci ledviny pomocí metod strojového učení

Autor: Kyrylo Stadniuk

Obor: Aplikovaná Informatika

Druh práce: Bakalářská práce

Vedoucí práce: Ing. Tomás Kourim Mild Blue, s.r.o., Plzenská 27, Praha 5

Konzultant: Ing. Pavel Strachota, Ph.D. Katedra matematiky, Fakulta jaderna a fyzikálne inzenyrska, Ceské vysoké udeni technické v Praze, Trojanova 13, 120 00 Praha 2

Abstrakt: Abstrakt max. na 10 řádků.

Klíčová slova: klíčová slova (nebo výrazy) seřazená podle abecedy a oddělená čárkou

Title:

Estimating patient's life expectancy after a successful kidney transplant using machine learning methods

Author: Kyrylo Stadniuk

Abstract: Max. 10 lines of English abstract text.

Key words: keywords in alphabetical order separated by commas

Contents

Chapter 1

Introduction

The goal of this paper is to explore fields of kidney transplantation and machine learning, create and apply machine learning model in real-world application.

Chapter 2

Medical Background

2.1 Why kidney fail

2.2 The history of kidney transplantation.

2.2.1 Early animal experiments

Advancements in surgical methods and techniques at the beginning of the 20th century eventually led to experiments with organ transplantation. On March 1st, 1902, Emerich Ullman, a physician at the Vienna Medical School, performed the first recorded organ transplantation. He performed an autograft, meaning the transplantation where the donor and the recipient are the same individual. Ullmann utilized the method of vascular suturing developed by Ervin Payr, to connect the dog's kidney to the vessels of its neck. The transplant was successful - the kidney produced urine. The dog was presented the same day to Vienna medical society eliciting significant interest and discussion.

The same year other similar transplantations were made. Another physician, Alfred von Decastello, performed a dog-to-dog kidney allograft at the Institute of Experimental Pathology in Vienna. The kidney produced urine for a while but then stopped working. Later Ullman performed a dog-to-goat kidney xenograft (cross-species transplant), and to his surprise kidney produced some urine, but later stopped.

In Lyon in the department headed by Mathieu Jabourday, his assistants Carrel, Briau and Villard were working on new methods of vascular suturing. In 1902, Alex Carrel published the method of vessel anastomosis now referred to as Carrel's seam. This technique represented a significant improvement over existing methods and effectively addressed the common issues of thrombosis, hemorrhage, stricture, and embolism[13].

Later Carrel moved to the United States where he continued his research on vessel suturing and organ transplantations at The Rockefeller Institute for Medical Research. There he perfected his method and while performing autografts and allografts documented what later would be recognized as "rejection". For his works in 1912, he got the Nobel Prize in Medicine. By this time, his method of suturing had been widely adopted in human surgeries[6].

2.2.2 Early human transplantation

The first recorded human renal xenograft was performed by Mathieu Jaboulay in 1906. He chose a pig and a goat as donor animals and performed two xenografts. One kidney was transported to the arm and the second to the thigh. Each kidney functioned for one hour[14, 4].

The second and third human transplants performed by Ernst Unger were far more known. On December 10, 1909, he performed a kidney transplant from a stillborn baby to a baboon. Even though the kidney produced no urine, the postmortem showed that vascular anastomosis (connection of vessels) was performed successfully. This inspired Unger to perform another transplantation that same month, but this time monkey-to-human xenograft. The kidney was transplanted from an ape to dying from renal failure young woman. The kidney never worked.

These early experiments demonstrated that technically kidney transplantation was possible, but the mechanism of rejection was not yet fully understood. Carrel in his famous lecture about the future of transplantation (1914) to the International Surgical Society mentioned that the works of his colleague at the Rockefeller center J.B. Murphy might seriously impact the development of the field. Murphy found that irradiation and benzol treatment increased the graft survival of cancer in mice. This observation inspired Carrel to conduct his own experiments, wherein he irradiated recipients and found prolonged graft survival, but these experiments were never formally published[11].

The period of the 1930s and 1940s was rather stagnant compared to the beginning of the century. European surgical centers that studied transplantology before were in decline. Moyo Clinic in the US was conducting some cautious experiments without considering Carrol's works and attempts at immunosuppression.[13] However, there was a notable event during this period - the first human-to-human transplantation. It was performed by Yurii Voronyi (in literature for some reason he is referred to as Voronoy) on March 3, 1933, in Kherson, Ukraine. The recipient was a 26-year-old woman admitted to the hospital on March 3, 1933, with mercury chloride poisoning induced by a suicide attempt the previous day that resulted in acute renal failure. Transplantation seemed the only viable option. It was known from previous experiments by other scientists that no xenograft ever was successful so human-to-human transplantation was the only feasible choice. The option of injuring a living person by organ removal was not even considered. It was known from the physiology that kidneys save their function a couple of hours after the reperfusion with ringer-solution and that organs keep some sterility a couple of hours after the host's death. So temporary cadaver transplantation until the woman's own kidneys would regenerate seemed to be a reasonable option. The transplantation was performed to the thigh's artery and vein using Carrel's seam with some modifications. After some time the kidney started to produce urine for a while but then eventually the allograft failed and 48 hours after the surgery the patient dies. The reason for graft failure was blood group mismatch and too long warm ischemia time - 6 hours, so the kidney began to degrade, resulting in an immune reaction to dead kidney cells and kidney blood cells. Voronyi performed another 5 such transplantations, which he considered as a bridge therapy until the recipient's own kidneys would recover. Kidneys produced urine for different durations from 1 to 7 days with 2 patients eventually recovering and living normally thereafter[12].

2.2.3 First successes

In 1946, at the Peter Bent Brigham Hospital in Boston, a group of surgeons: Hufnagel, Hume, and Landsteinerhuman performed kidney transplantation under local anesthetic on the arm vessels. The short period of kidney functioning may have helped the patient to recover from acute renal failure. It ignited the hospital's interest in renal transplantation.

Simonsen in Denmark, Dempster in London, and Küss in Paris concluded that it is preferable to place the kidney in the pelvis. Further, both Simonsen and Dempster deduced that the immune response was responsible for graft failure and both hypothesized that the humoral mechanism of rejection was probable.

In the early 1950s, two groups of surgeons based in Paris and Chicago performed pelvic kidney transplants without immunosuppression. In Paris, Jean Hamburger reported the first live-related kidney

transplant between a mother and her child. he transplanted kidney began to function immediately. It functioned for 22 days until it was rejected.

A series of nine transplantations with the thigh position of the allograft was closely studied in Boston and the first usage of hemodialysis for the preparation was recorded in Boston by David Hume in 1953. In some of these cases, mild successes were achieved using the adrenocorticotropic hormone (more known as cortisone). It was hypothesized that the endogenous immunosuppression of uremia was responsible for the results rather than the drug regimen. Hume's findings were substantial as he concluded that prior blood transfusions, blood group matching between the donor and recipient, and host bilateral nephrectomy could be beneficial for the success of the transplant. These conclusions were later confirmed by subsequent studies.

These attempts in the early 1950s taught technical aspects of kidney transplantation and with increased confidence on December 23, 1954 in Boston Joseph Murray performed kidney allograft from one identical twin to another, bypassing the rejection barrier. From that time many similar surgeries were performed in Boston. This caused a lot of talks and predictions but all of them were negated when one of such recipients got pregnant and gave birth to a completely normal infant. However, in retrospective, it didn't bring anything new scientifically, because the technical possibilty of kidney transplantation was evident and the cases of successful skin allografts between identical twins were known for decades, but nonetheless it was an important milestone that aroused the interest in further experiments [14, 4].

2.2.4 Attempts in immunosuppression

In 1948 at Mayo Clinic patients handicapped by rheumatoid arthrytis were given already mentioned cortisone, adrenal cortical hormone with mild immunosuppressive properties, that relieved their condition. This popularized the research on adrenal cortical hormones, but later it was concluded that the steroid effect was clinically insignifficant for transplantation. After that, the experiments with irradiation, abandoned by Carrel and Murphy, were revitalized. Joan Main and Richmond Prehn showed that weakening of the immune system of adult mice by radiation and consequent skin and bone marrow transplantation from the same donor resulted in skin transplant acceptance. This encouraged teams in Boston and Paris to pursue the similar approach in humans.

In 1958, Murray's team transplantation on humans utilizing the Main-Prehn method conducted lethal total body irradiation (TBI) on two patients with additional bone marrow transplant. Ten more recipients were irradiated with sub-lethal TBI, but without donor bone marrow transplant. As a result 11 patients passed away within a month, the only survivor had sub-lethal TBI without transplanted bone marrow and he got kidney from his non-identical twin brother. This was rather revolutionary - for the first time kidney was not rejected from non-identical twin. The kidney functioned for 20 years. Jean Hamburger and his team performed another fraternal twin transplant utilizing the same irradiation technique. The transplant functioned for 26 years finishing with the recipient's death for rejection-unrelated reason.

Between 1960 and 1962 Kuss and Hamburger performed four successful transplantations between non-twin patients with following TBI. This gave promise that the transplantations could be done in non-twins and potentially between anybody. The research continued.

It was obvious that TBI is not the best choice and that it is necessary to find a substitution. In 1959, Schwarz and Damesehek from Tufts University published paper that described how an anticancer drug 6-mercaptopurine (6-MP) lowered immune response to foreign proteins in rabbits. Roy Calne, a training surgeon at Royal Free Hospital, London, dissapointed with TBI in prolongation of kidney allograft survival in dogs, noticed Schwarz and Dameshek's paper and performed his own experiment in dogs and found that it signifficantly prolonged dog's survival. Charles Zukoski and David Hume found the same outcomes.

6-MP was used in three tranplantations at Royal Free Hospital, but without success, However Kuss and assosiates reported one prolonged graft survival from a nonrelated donor. The TBI was main agent and intermittent usage of azathioprine and prednisone was used as an additional therapy.

Gertrude Elion and George Hitchings provided Roy Calne with the 6-MP derivative - azathioprine. Calne showed even longer graft survival with azathioprine. Both Elion and Hitchings were awarded with the Nobel Prize for the development of 6-MP and azathioprine. In 1961 Azathioprine became available for human use.

2.2.5 Gloom then revolution

In 1963 National Research Council organized a small conference constisting of 25 transplant clinitians and scientists to review the status of kidney transplantation at the moment.'the discussion was quite depresive. Clinitians presented their results, that were rather discouraging: less that 10% of hundreds performed transplantations survived for more than three months, from patients with TBI only six got to the one year mark. Murray reported that from his first ten patients on 6-MP one survived for a year, others passed away within 6 months, so it was concluded that drugs were not more effective than radiation.

The gloom continued untill Tom Starzl, until then unknown, did his presentation where he described his protocol that allowed graft survival for more than one year in 70% of cases. He was not believed at first, but then he showed medical records of his patients and he was eventually believed. The only thing that differed from other protocols with 6-MP was that addition of prednison. This was a sensation. In the first year after the presentation, 50 new tranplantation programs were founded in US alone. And his protocol became medical world standard for the next 20 years.

2.2.6 Plateau

During the period from 1964 to 1980 nothing groundbreaking had happened, although the steady development was seen. Dialysis become available and thanks to the acumulated exprerience the dosages became more precise. The brain death was accepted and the body was supported for a while to save organs for transplantation.

Hemodialysis for renal failure was created by Willem Kolff from Holland during WWII. But it couldn't be used for chronic renal failure untill 1960 when was invented Teflon arteriovenous conduits for long-term vascular access.

Acceptance of brain death as a real death. Before the mid 60s the kadaver transplantation was limited by the ischemic damage. Now the additional organs were available from "heartbeating kadavers".

Cold for organ preservation. This was suggested in 1905 by Carrel's colleague Charles Guthrerie. Initially, Starzl used total body hypothermia to protect donor organs, but by 1960 switched to infusing cold solution into the portal vein to protect donor livers. In 1963 the infusion of cold solution intravenous in the transplanted kidney has become a standard.

As the organ preservation for more than 6 hours was achieved in mid 60s the exchange of organs between centers has became practical. Initially sharing was local and informal, that roused the worry that the organs could be distributed unequally and that they could be transported outside of the US. This led to Congress passing the National Transplant Act in 1984. The Southeastern Organ Procurement Foundation (SEOPF), founded in 1969 and eventually composed of 12 hospitals in several cities, served as the template for the United Network of Organ Sharing (UNOS) that controls organ allocation and placement, monitors performance of transplant centers and organ procurement organizations, collects data, and controls quality. They kindly provided us with data for this paper.

2.2.7 Tissue typing

Although tissue typing was suggested by Alexis Carrel in the beginning of 20th century it could not be proven and used until 1958 when Jean Dausset discovered the first human leukocyte antigen (HLA). Testing for antibodies was not reliable until 1964 when Paul Terasaki invented a microxytotoxicity assay. Test included mixing donor's lymphocytes and recipient's serum and quickly has became the standard and was name crossmatch. For a couple of years Terasaki performed typing for most of U.S. transplant centers and found a couple of observations: 1) Positive cross-match test predicts hyperacute rejection. 2) matching can reliably identify optimal donor within a family. It was assumed that the same would work for non-related recipients.

However, when in 1970 Terasaki reviewed his large database of cadaver renal allografts he found no correlation with the typing. This raised a lot of agitation in tissue typing community and his grant even was temporalily suspended until others didn't report the same. Now it is conclded that the

2.2.8 Antilymphocyte serum

Next mark was cyclosporine, a fungal derivative with immunosuppressive properties discovered in 1976 by Jean-François Borel. It revolutionized the renal and extrarenal transplants, proving to be much better than the previous drug azathioprine. However it also had to be combined with prednisone to gain those results. It was used until 1989 when even more potent drug was discovered - Tacrolimus. It helps even when the cyclosporine with prednisone has no effect.

Tom Starzl discovered that donor leukocyte chimerism was present in patients who had maintained successful kidney or liver grafts for up to three decades.

chimerism is an important cause (not the consequence) of successful transplantation,

successful engraftment is the result of the responses of coexisting donor and recipient cells each to the other causing reciprocal clonal exhaustion followed by peripheral clonal deletion

2.2.9 Conclusion and challenges of the field

The ultimate goal is immunosuppression without drugs because drugs are often toxic and the proper dosing might be tricky.

2.3 Immunology

The immune system is a sophisticated defense mechanism that evolved to protect multicellular organisms from pathogens such as bacteria, fungi, viruses, and parasites. It consists of many cells and tissues that compose a complex system that detects, evaluates, and responds to the invader. The immune system is divided into humoral and cell-mediated immunity. Humoral is mediated by soluble immunoglobulin proteins referred to as antibodies, while cell-mediated involves pathogen-specific T Lymphocytes that either destroy the invader or assist other cells in doing so. Both are essential for a complete immune response.

Lymphocyte is a type of white blood cell that is responsible for both humoral and cell-mediated immune responses. There are two types of lymphocytes: T lymphocytes (T cells) and B lymphocytes (B cells). B cells mediate humoral response by producing antigen-specific antibodies. An antigen is any molecular structure that binds to an antibody or specific surface T cell receptor, triggering an immune response. Once B-cell encountered an antigen it starts to produce antibodies specific to it, antibodies then bind to it, marking the invader for destruction. T cells when encountering an antigen start to proliferate

forming an army of T cells that will eliminate the invader and will form long-term memory about the pathogen.

Physical barriers: epithelia and mucous membranes constitute the first line of defense. To activate the immune system the pathogen must first breach physical barriers. The immune system categorizes pathogens by common characteristics and designs its response accordingly. Pathogen detection and categorization rely on the interaction between pathogen and T-cell receptors, as well as soluble antibodies. Binders for T cell receptors and antibodies can be the whole pathogen's body, its part, or molecules excreted by it.

Pathogens are recognized and categorized by molecular patterns that are associated with a particular pathogen and are referred to as pathogen-associated molecular patterns (PAMPs). Pathogen recognition receptors (PRRs), which are excreted by white blood cells, bind to PAMPs initiating the cascade of events that will mark a pathogen for destruction.

Pathogen-host interaction is a continuous arms race, as pathogens usually have a short life cycle and can modify their DNA to elude the host's recognition systems. The generation of diversity in developing cells is designed to combat this. When lymphocytes are developing in bone marrow random PRRs are generated, then cells are tested on non-reactivity to host cells. If the test is passed the cell is released into circulation. The principle of recognizing self vs. non-self is called tolerance.

There are two interconnected systems of response: innate and adaptive. Innate includes primitive built-in cellular and molecular mechanisms aimed at preventing infections and quickly demolishing common pathogens. It consists of physical and molecular barriers as well as PRRs that are encoded in DNA and therefore are inherited. Innate immunity provides a fast and effective response which however is not very specific and cannot differentiate small differences. Adaptive immunity is constituted by both humoral, where antibodies neutralize and eradicate extracellular microbes and toxins, and cell-mediated immunity, where T lymphocytes exterminate intracellular invaders.

Adaptive immunity is much slower but more able to recognize small differences. It typically starts to act within 5 to 6 days after initial exposure. Because it takes time to create an army of cells with specific receptors. After pathogen extermination, some of the lymphocytes with the specific receptor become memory cells, making it easier to fight this type of pathogen.[13]

In conclusion, the immune system is a complex network of molecules, cells, tissues, and organs that cooperate in protecting the organism from pathogens. The system can be divided into two main branches: innate and adaptive, which cooperate in protecting the host from infections while developing long-term immunity to specific pathogens. Understanding the mechanisms of the immune system is essential to understanding the domain of kidney transplantation.

2.4 Immunology of kidney transplant

The process of transplantation inevitably includes termination of blood flow, and, as a result, oxygenation. Therefore cell is unable to generate sufficient amount of energy to maintain homeostasis, leading to damage or death. Damage or death is associated with DAMP release that might be detected by both innate and adaptive immunity.

2.4.1 Immune system activation Peritransplant

The process of transplantation inevitably includes termination of blood flow, and, as a result, oxygenation. Therefore cell is unable to generate sufficient amount of energy to maintain homeostasis, leading to damage or death. Damage or death is associated with DAMP release that might be detected

by both innate and adaptive immunity. Mostly it is the ancient innate immunity that is activated with its soluble arm - complement system.

Damage signals Many DAMPS are recognized by the same PRRs that mediate response to PAMPs. These DAMPS include molecules that are normally hidden from the immune system and are produced during ischemia, such as extrecellular ATP, heat shock proteins (HSPs), uric acid, etc. Likewise, oxidative stress and decline in intracellular potassium may act as intracellular damage signals.

Complement Complement system is comprised of series of proteine kinases that are sequentially activated resulting in membrane attack complex (MAC) formation. MAC include complement components C5 to C9, which are inserted into pathogen cell membrane resulting in compromising cell integrity leading to cell death.

There are three pathways of complement system activation: the classical pathway, the alternative pathway, and the mannose-binding lectin (MBL) pathway. The classical pathway is activated by IgM and IgG antibodies and participates in antibody-mediated rejection, that will be discussed further. Alternative complement is always active and therefore must be controlled by a regulatory proteins, to prevent inadequate responses. The MBL pathway is activated by damaged endothelium, a cell tissue that covers organs and vessels, and carbohydrates present on pathogens. Either pathway results in C3 convertase that cleaves C3. This cleavage leads to a cascade of reactions that culminate in MAC formation.

Long ischemia time results in endothelial cell damage that is acossiated with ischemia-reperfusion injury (IRI). IRI activates MBL and alternative complement pathways.

Gene silencing using small interfering RNA (siRNA) might be a promising instrument in organ transplantation, because it can be appliend to an allograft during cold reperfusion and it has been shown to mitigate IRI in animal models. Other strategies of supressing local complement activation would also be useful.

2.4.2 Stimulation of Adaptive Alloimmunity

Immune response to a graft occurs in two main stages: afferent and efferent arms. In afferent stage, recipient lymphocytes are stimulated by donor antigens and start to proliterate and send signals to other cells. In efferent arm, leukocytes migrate to the transplanted organ and donor specific antibodies are produced.

For the immune system to be activated graft must express antigens that will be considered by the host's immune system as foreign. These include ABO antigens, human leukocyte antigens (HLA), and polymorphic non-HLA "auto-antigens".

ABO blood group antigens

ABO system is used to group blood into groups, based on presence or absense of antigens on a blood cell surface. There are four major blood groups: A, B, O and AB.[7]

When allocating an organ to transplant the first thing that is considered is ABO blood group antigens compatibility. ABO antigens are expressed almost by any cell in the allograft, and if the transplantation to be carried out in ABO-incompatible donor and recipient it would result in a hyperacute antibody-mediated rejection.

Donors with blood group O are so called "universal donors". Organs from them can be safely transplanted to recipients with any ABO blood group. Whereas, recipient with AB group can safely receive organ from recipient with any ABO blood group and is called a "universal recipient".[9]

Table 2.1: MHC class division

MHC class I	MHC class II
HLA-A	HLA-DR
HLA-B	HLA-DP
HLA-C	HLA-DQ

HLA

Histocompatibility antigens are genetically encoded antigens that cover cell surfaces. They differ between individuals of the same species and therefore trigger an immune response in case of allograft. In all vertebraes histocompatibility antigens are divided into single major histocompatibility complex (MHC) and numerous minor histocompatibility (miH) systems. In case of either MHC or miH incompatibility the result is an immune response to the graft, more severe in case of MHC than miH. Rejection in MHC-compatible donor-recipient pair is usually delayed, in some cases forever. Although, sometimes miH mismatch might be so severe that it would be comparable to full MHC mismatch.

MHC antigens are proteins that cover cell surfaces to help the immune system to recognize self vs. non-self. Major histocompatibility comples is divided into MHC class I and MHC class II. MHC class I cover surfaces of most cells and are liable for activation of cytotoxic CD8 cells, that help to find and destroy infected cells. MHC class II are found on certain immune cells and play crucial role in immune response coordination. In humans MHC class 1 are divided into three subgroups each, as can be seen on table

In clinical practice, clinicians asses and try to match donors and recipient according to the number of HLA-A, -B, and -DR mismatches, ranging from zero mismatches (0-0-0) to a maximum of 6 mismatches (2-2-2). Generally more emphasis is placed on DR loci due to capability of CD4 T cell activation, which might trigger both humoral and cellular adaptive immune responses.

Minor histocompatibility proteins can act as antigens, although weaker than MHC. However if prior sensitisation exists it could result in severe immune response that might result in graft loss.

2.4.3 T Cell-mediated rejection

T cell-mediated rejection or TCMR is the most common type of allograft rejection, as it still happens in 20% of transplantations mostly within first 6 months posttrasplant. Immune system cells migrate through vessels to the graft, become activated and start to attack the organ. Complement may also play role in it.

2.4.4 B Cell-mediated rejection

B cells are immune system cells that produce antibodies. Alloantibodies are antibodies that react to donor-specific HLA antigens and might cause hyperacute rejection, acute antibody-mediated rejection (ABMR), and chronic ABMR. About 30% of patients have sensitivities and have certain HLA antibodies. It might cease transplantation or require antibody suppression strategy. Even low amount of antibodies below crossmatch cutoff doubles the risk of ABMR and increases the risk of graft failure by 76%. Additionally, donor specific antibodies might develop posttransplant and cause an actue ABMR.

Acute AMBR is rarely seen in patients withour prior sensitization and is highly difficult to treat. AMBR is characterized by decline in allograft function, presense of DSA and signs of acute vascular injury. A progressive reduction in graft function over time is observed almost universally.

2.4.5 Transplant tolerance

Taking into account the detrimental effect of long-term immunosuppression one of the primary objectives in transplantation is the induction of immunologic non-responsiveness (tolerance) to an allograft. The are a couple a pathways of immune non-responsivenes generation described in literature, however it haven't gone further animal models yet.

2.4.6 Factors Influencing rejection beyond the graft - microbiome

Human body is very complex system where every subsystem influences other subsystems and the whole system in general. It is clear that gut microbiome has a profound influence on immune system. It is possible microflora on the allograft might cause rejection. Immunosupression, profilactory antibbiotics, diet changes and other restrictions asossiated with organ transplantation result in decrease in gut microbiome diversity that result in systematic inflammation, that might contribute to alloimunnity, as well as autoimunnity.

2.5 Conclusion

Chapter 3

Machine Learning Background

What is ML: Machine learning (ML) is a subfield of computer science that consists of building algorithms that are capable of processing large amounts of data, finding patterns and making some actions such as predictions or even generating new data. Machine learning is an intersection of many fields of science such as statistics, theory of probability, linear algebra, calculus and, of cource, computer science. As Arthur Samuel eloquently put it: "[Machine learning is the] field of study that gives computers the ability to learn without being explicitly programmed."

Why is this important: Machine learning excels in problems that are either overly complex or have no known algorithm.[9] It can help us learn. We can extract previously unknown correclations from the data and create knowledge. It might have less error in decision making than humans.

How it is divided: Based on the problem and, therefore, on our approach to building the dataset(tell somewhere the difference between the dataset and raw data) and the model, machine learning can be divided into four sub-fields: supervised, semi-supervised, unsupervised and reinforcement learning. Supervised learning means that we labeled our data and we want to predict labels from the data. (explain labeled data!) Semi-supervised learning means that only a part of our data is labeled. Unsupervised means that our data is unlabeled. In reinforcement learning we create an environment, set up rewards for doing certain things and punishment for other things, and let the machine (actor) to perform actions that produce the highest reward. A separate subsection is dedicated to each subfield.

How it is important in medicine: Every field suffers from human errors, and medicine is no exception. (people get tired, wrong institutions, etc...) According to Makary et. al., medical error is the third death cause in the United States [11]. Machine learning also make mistakes, but if we manage to get at least 1% less error than human error this will be a success. Human body is a complex system, where other humans (doctors) can't possibly comprehend everything that is going on and how it is related to each other. Also it can help us gain insights from data that were already accumulated and possibly make some discoveries.

Why machine learning in not learning: It is important to note that machine learning is rather marketing term. Machines can't learn. If we were to at least slightly distort inputs, machine wouldn't be able to "learn". This is very different from learning in humans and animals.[10] It is debatable that statistics and mathematics can really produce consciousness. Luckily the debate lies beyond this paper.

In this chapter we will cover all theoretical background needed for solving our problem, this includes classical machine learning, deep learning, statistical survival analysis, basic steps that are required to create machine learining systems, how to preprocess data and what algorithms are worth of our attention in solving tasks like this. We will begin by exploring what is supervised learning.

3.1 Supervised Learning

Supervised learning is the process of training a model on data where the outcome is known to make predictions for data where the outcome is not known[12]. *Classification* and *regression* are common supervised learning tasks. In this section we will define these problems and necessary terminology, describe commonly used algorithms that are used to solve these types of problems.

In supervised learning the *dataset* is the collection of labeled examples $\{(x_i, y_i)\}_{i=1}^{N}$, where each individual x_i is called a *feature vector*. A feature vector is a vector that in each its dimension $j = 1, ..., D$ contains a value that describes an example in some way. This value is called a *feature* and is denoted as $x^{(j)}$. The *label* y^i might be either a finite set of classes $\{1, 2, ..., C\}$, in case of classification task, or a real number, a vector, a matrix or graph, in case of a regression. The goal of supervised learning algorithm is to create a model using the dataset that will take the feature vector as input and produce label or more complex structre as a output.

Classification is a problem of assigning a label to unlabeled example. This problem is solved by a classification learning algorithm that takes a labeled set of examples as an input and produces a model that takes unlabeled example as input and outputs a label, a number associated with it or a probability of belonging to a certain class out of which it is easy to deduce the class. If the set of labels has only two classes we talk about *binary classification*. Consequentlhy, if the set of labels has three or more classes it is a *multiclass classification*. Some algorithms are binary classifiers by definition, while others are multiclass classifiers. It is possible to create an *ensemble* out of binary classifiers that will be able to perform multiclass classification. Ensemble is a combination of algorithms that are connected together to perform one task.

Regression is a problem of predicting a *target value* given an unlabeled example. The problem is solved by regression learning algorithm that takes a set of labeled examples as inputs and produces a model that takes unlabeled example as input and outputs a target value.

3.1.1 LinearRegression

Linear regression is a popular regression learning algorithm. The model it produces is a linear combination of all features

What is the model:

What we are trying to minimize(loss function):

3.1.2 LogisticRegression

probability of an instance belonging to a certain class.

3.1.3 SVM

kernels
support vectors

3.1.4 DecisionTrees

highly interpretable model
builds trees that can be visualized and show the model's reasoning

3.1.5 RandomForests

The concept of an ensemble:
Stacking:
(some other technique)
Base learner of the random forest:
How it all works together:
Benefits: potentially interpretable model.

3.2 Unsupervised Learning

Unsupervised learning is the kind of learning that operates with the dataset without lables. It proves useful during EDA and *dimensionality reduction*, but otherwise it is not very useful in practical applications, as unlabeled dataset means no feedback about how well your model performs.

We will cover this part very briefly, as it is not closely related to our task. Albeit something will be used as a data preprocessing step, such as dimensionality reduction, or during exploratory data analysis.

There are more approaches but we are going to cover only three, that might prove useful to our task. *Clustering* is identifying similar instances and assigning them to groups of similar instances - clusters. It can be used for data analysis, customer segmentation, as a dimensionality reduction technique, anomaly detection. Clustering might be *soft* and *hard.* Hard clustering means that an instance might belong to only one class. In soft clustering instance has a score of belonging to a particular cluster. The score might be the distance from the cluster centroid or an affinity (similarity score).

Dimensionality reduction is useful for visualization and for acceleration of learning. Datasets often have a lot of redundant data or it might be the case that the task requires a lot of features. A lot of algorithms, such as linear model, SVMs, decision trees, do not handle high dimensional data well. So called *curse of dimensionality* states that high dimensional data can cause slow learning and prevent us from getting an optimal model. So the reduction of the data dimensionality might be a good idea. Its worth noting though that dimensionality reduction algorithm might loose some useful information, while decreasing training time. A lot of modern algorithms, such as neural networks or ensemle algorithms, handle high dimensional data very well, so dimensionality reduction techniques are used less than in the past. They are still used however for data visualisation and for cases when we need to build an interpretable model and we are limited in number of algorithms we can use.

Anomaly (*outlier*) *detection* involves detection of instances strongly deviating from the norm. These instances are called *outliers* or *anomalies*, normal instances are referred to as *inliers.* It is useful in many applications. It can be used as a data preprocessing step - to remove outliers from the dataset, which might improve the performance of the resulting model. Also, it can be used in *fraud detection* task and in detection of faulty products in manufacturing facility.

*Novelty detectio*n is closely related task to the anomaly detection. The only thing different about them is that the novelty detection makes an assumption that the dataset the model was trained on was not contaminated by outliers, while the anomaly detection does not make this assumption.

3.2.1 KMeans

Finding optimal numbers of clusters:
Advantages of kmeans
Limitations of KMeans:

3.2.2 Principal Component Analysis (PCA)

Principal components are vectors that define new coordinate system. First vector goes in the direction of the highest variance. Second vector is orthogonal to the first one and goes in the direction of the second highest variance, and so on. If we were to reduce dimensionality to $D_{new} < D$, we would pick D_{new} largest principal components and *project* instances onto them.

(Create images)

It is not adviced to choose the number of dimensions arbitrarily. Usually such number of dimensions is chosen that preserves large amount of variance (e.g. 95%), or in case of visualization we reduce number of dimensions down to 2 or 3.

There are different versions of PCA; kernel PCA, Incremental PCA (online or batch PCA), Randomized PCA. But covering those lies beyond this paper.

3.2.3 Gaussian Mixtures

Gaussian mixtures is common algorithm that can be used for anomaly detection. Gaussian Mixtures assume that the dataset is generated by several Gaussian distributions. Any instance lying in the region of low density is an anomaly. The density threshold has be specified. If one gets too many *false positives* (good products labeles as faulty) them need to decrease the the threshold, consequently if we get too many *false negatives (*faulty products labeled as good) the threshold has to be increased. Gaussian mixtures belong to *soft clustering.*

Gaussian mixtures require the number of clusters to be specified. It needs to be run a couple of times to avoid suboptimal solutions.

3.3 Semi-supervised learning

3.4 Reinforcement Learning

mention RLHF - reinforcement learning through human feedback. LLMs

3.5 Data Preprocessing

Feature engineering: Feature engineering is the process of creating new features from the raw data. For example, calculation of eGFR, metric of kidney function estimated on patient's age, gender and serum creatinine level, will potentially give more information to the learning algorithm, than all those features separately.

- **One hot encoding**: One-hot vector is vector representation of

One hot encoding is used for categorical values, such as gender, age group, etc, where there are not too many potential values.

- Ordinal encoding
- Binning

Normalization:

Standartization:

- **Dealing with missing features**: There are a couple of popular ways of dealing with the missing values. Here are a couple of them:

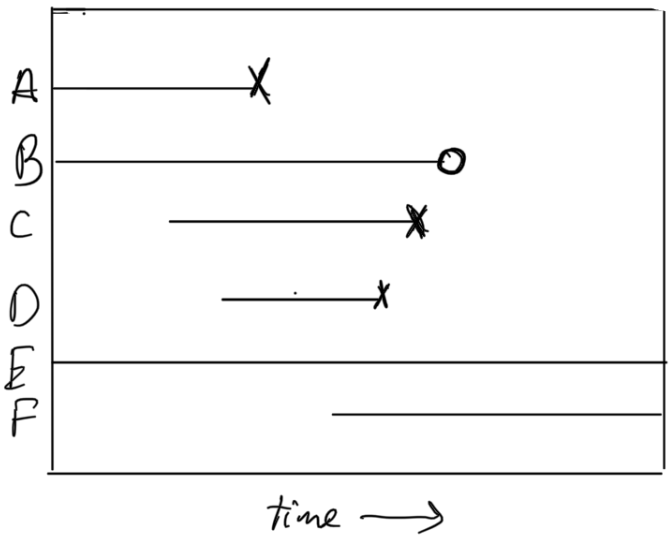


Figure 3.1: caption

– dropping rows with missing values. The most obvious and bruteforce way of dealing with them. If there are not so much of missing values, or we have a dataset that is large enough, the usage of this technique would be appropriate.

– filling the missing values with median, or clearly too high or too low value for the algorithm to understand that it is an outlier and this particular value should not be considered, while other features might be.

Data imputation:

-

3.6 Survival Analysis

Definition: *Survival analysis,* also known as *time-to-event analysis*, is a statistical method used to analyze the time until an event of interest occurs. Its name originates from the clinical and biological research, where these methods are used mostly to analyze survival time, hence the name. These methods, however, found their uses in areas far beyond clinical setting: in business to analyze the time until the customer ”churns” from a subscription, in engineering, to estimate the longevity of certain products, or its parts. In social sciences, estimate the longevity of a marriage or a dropout rate in academic setting.

Censoring: The most pronounced feature of these methods is the ability to handle censored data. Censoring is a situation when the information about the survival time is known only partially. For example, in the dataset used in this work there are **(number of patients)** that were still alive until last date of observation **(date of observation**). And we do not know what happened to them after that date — they are censored.

Look at the figure 3.1. On the y axis we can see individual patients. x axis corresponds to timeline of the study (right side is the end of the study. Cross (X) denotes an occurance of the event, circle (O) corresponds to subject’s exit from the study.

Left and right censoring: There are two types of censoring: left and right censoring. Right censoring, the most common one, happens when we know that event didn’t happen up to a certain point — the

patient didn't attend follow-ups after a certain time or the study ended when he was still alive. Lets look at the figure 3.13.1, E and F are obviously right censored, as the event didn't happen during study and B is also right censored, as the subject dropped out of study. Left censoring, being much less common, describes the situation when the subject's survival time is unkonown, but we know that it is less than a certain time. For example, wh (**add appropriate left censoring example**)

Censoring assumptions: They are very similar, however, they have very subtle differences that we are going cover in the following paragraphs.

-Random: censored subjects are representative of the remaining subjects provided the same survival experience.

-Independent: censoring is independent if it is random within any subgroup

Random an independent censoring assume that those who left the study are no different from those who stayed and therefore the percentage of survival is the same in both groups with respect to their survival experience. They are the same if assumed in one group. Censoring is *independent* if the censoring percentage in one group differs from the censoring percentage in the other group. In contrast, if the censoring percentages are the same, the censoring is *random*.

-Non-informative censoring: distribution of survival times T provides no information about the distribution of censorship times C.

Why survival analysis is important :

Where is it used :

Relation to machine learning: Because we are predicting a numerical value(rather values), from the machine learning point of view, survival analysis might be considered a regression, but instead of one numerical value we predict a continuous value - survival function or hazard function, and, obviously, the dataset is censored. *Survival function (also survivor function)* $S(t)$ shows us the probability of patient *surviving (event doesn't happen)* at a given time t and can be denoted as

$$S(t) = P(T > t). \tag{3.1}$$

Where t is any specific *time* of interest, T is random variable for subject's survival time. For instance, if we want to know if a patient is going to live for more than 5 years after kidney transplant, t is equal to 5 and we ask whether T is greater than t (probability question). The function is declining in the range from 0 to infinity. As it is a probability, the function value ranges only from 0 to 1. Theoretically, the graph of the survival function must be smooth, but in reality it is represented by a step function.

Hazard function $h(t)$ tells us the probability of given event *happening* at a given point of time t, provided the event did not happen before time t, and is denoted as

$$h(t) = \lim_{\Delta t \to o} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \tag{3.2}$$

Subject's survival time T lies between t and $t + \Delta t$ provided that survival time T is greater or equal than t. Sometimes, the hazard function is called a *conditional failure rate.* It is a rate because it is a conditional probability per unit of time Δt. As it is not a probability, but a rate, the scale for this ratio is from 0 to infinity — depends on the measure of time in days, weeks or years. When we consider the limit of the expression as the time interval approaches zero we basically get the instantenious potential of failing at time t per unit time, given survival up to time t.

Cumulative hazard function:

The relationsip between the two: There is a clear relationship between the survival function $S(t)$ and the hazard function $h(t)$ – if we know one, we can determine the other. The relationships are the following:

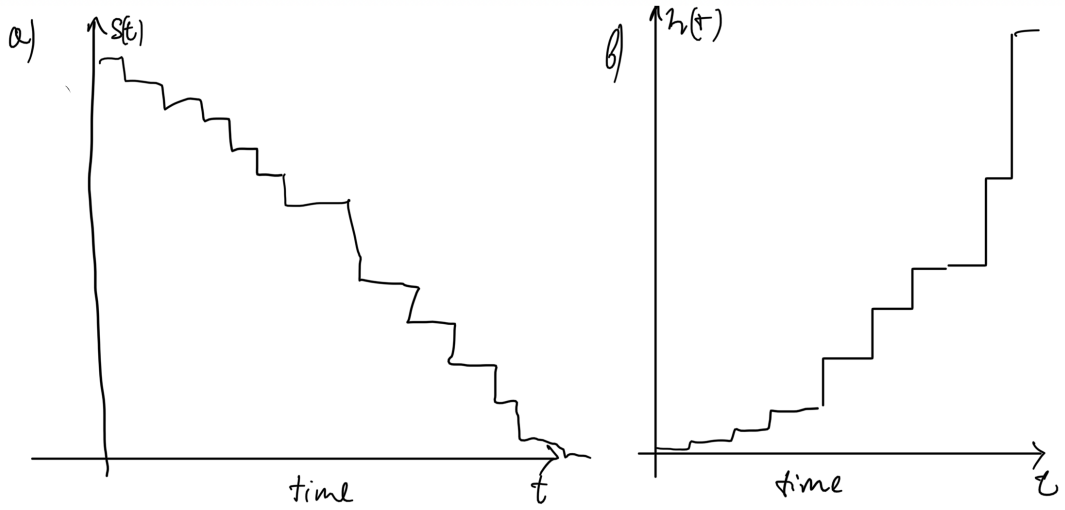


Figure 3.2: (Don't forget to provide example images both for the survival and hazard functions for some random patient from the dataset)

$$S(t) = exp\left[-\int_0^t h(u)du\right] \tag{3.3}$$

Equation ... tells that the survival function $S(t)$ is equal to the exponential of the negative integral of the hazard function from zero to t.

$$h(t) = -\left[\frac{dS(t)/dt}{S(t)}\right] \tag{3.4}$$

Equation ... tells us that the hazard function is equal to the negative derivative of the survival function $S(t)$ with respect to t divided by $S(t)$.

Considering the facts that survival function describes the probability of patient surviving to a given point of time t and hazard function shows us the probability of person dying at any given point of time t, we can say that they provide complementary information about survival and risk over time. Of the two discussed functions the survival function is used much more often as it is more appealing in the context of survival analysis and in the practical part of this paper(bachelor project/thesis) I am going to estimate exactly the survival function.

Take a look at figure 3.2. *a)* shows us a graph of the estimated survival function and b) shows us a graph of the estimated hazard function for a random patient from the dataset used. As we can see the survival function is declining over time, while the hazard function increases.

Kaplan-Meier Survival Curves:

it is one of the ways to create a survival function

Log-Rank Test: It is a way to compare two survival functions. Often used in studies, where there is a target group and a placebo (control) group to assess the efficacy of the studied thing by comparing the survival curves of the two groups.

Goals of survival analysis:

3.6.1 Performance Metrics

c_index: Begins to be biased at high levels of censoring
Uno's c_index: Handles high levels of censoring (provide a number) very well
Time-dependent Area under the ROC:
Time-dependent Brier Score:

3.6.2 Survival Gradient Boosting

—
Survival analysis methods
- CoxPH
- Regularized Cox
Cox Proportional hazards method - linear model
Survival Trees

3.6.3 Random Survival Forests

useless with large amounts of data

3.7 Deep Learning

neuron
neuron's parts
neuron activation function
layers
dropout
backpropagation
stochastic gradient descent
feed forward
different architectures
transformers
convolutional neural networks
natural language processing
neural networks for survival estimation

3.8 Machine learning workflow

0. define problem type and what objectives you want to achieve.
1. gather data
2. Analyse data
3. Create dataset: feature engineering
4. Apply algos for the given problem type
5. Choose the best, fine tune it, run on the whole dataset and/or modify the dataset
6. Repeat
7. Deploy

3.9 Challenges of machine learning

Poor quality data
Over

3.10 Overview of Machine Learning Libraries

3.10.1 Sci-kit learn

- pros
- cons
- where it is used

3.10.2 Keras

3.10.3 Tensorflow

- pros
- cons
- where it can be used

it is more corporate driven

3.10.4 PyTorch

- pros
- cons
- where it can be used

it is more research driven

3.10.5 Comparison

3.11 Conclusion

Chapter 4

Data Preparation and Analysis

4.1 Exploratory Data Analysis

- how many transplantations we have
- median age
- age distribution
- survival time distribution
- box plots

4.1.1 ...

4.2 Data preparation

4.2.1 Feature selection

4.2.2 Handling Numerical Values

4.2.3 Handling Categorical Values

Chapter 5

Machine Learning Model

5.1 Problem Definition

Clearly, the problem involves survival time and we have a censored dataset, therefore the survival analysis methods from the section 3.6 are the best choice

5.2 Model selection

bottleneck in a number of instances each model can train on. +

different groups of people have different expected life expectancy -> separate models for different groups of people for the result maximisation.

–

Create an ensemble of a couple of learners.

5.3 Model comparison

5.4 Final Model

5.5 Discussion

Chapter 6

Applications

6.1 Existing Solutions

6.1.1 Txmatching

6.2 KidneyLife

6.2.1 Frontend

6.2.2 Backend

6.2.3 MLOps

Conclusion

Text of the conclusion...

Bibliography

[1] Knechtle, S. J., Marson, L. P., & Morris, P. (2019). Kidney transplantation - principles and practice: Expert consult - online and print (8th ed.). Elsevier - Health Sciences Division

[2] Nobel prize in physiology or medicine (2022) Our Scientists. Available at: https://www.rockefeller.edu/our-scientists/alexis-carrel/2565-nobel-prize/ (Accessed: February 6, 2023).

[3] Barker, C. F., & Markmann, J. F. (2013). Historical Overview of Transplantation. Cold Spring Harbor Perspectives in Medicine, 3(4). https://doi.org/10.1101/cshperspect.a014977

[4] Matevossian, Edouard, et al. "Surgeon Yurii Voronoy (1895-1961)-a pioneer in the history of clinical transplantation: in memoriam at the 75th anniversary of the first human kidney transplantation." Transplant International 22.12 (2009): 1132.

[5] PUNT, Jenni et al. Kuby immunology. Eight. vyd. New York: Macmillan Education, 2019. ISBN 9781319114701;1319114709;

[6] ABBAS, Abul K., Andrew H. LICHTMAN a Shiv PILLAI. Basic immunology: functions and disorders of the immune system. Sixth. vyd. Philadelphia: Elsevier, 2020. ISBN 9780323549431;0323549438;

[7] NCI Dictionary of Cancer terms (no date) National Cancer Institute. Available at: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/abo-blood-group-system (Accessed: March 6, 2023).

[8] Dean L. Blood Groups and Red Cell Antigens [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2005. Chapter 2, Blood group antigens are surface markers on the red blood cell membrane. Available from: https://www.ncbi.nlm.nih.gov/books/NBK2264/

[9] Aurélien Géron. Hands-on Machine Learning with Scikit-Learn and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc., Sept. 2019.

[10] Andriy Burkov. THE HUNDRED-PAGE MACHINE LEARNING BOOK. Andriy Burkov, 2019.

[11] Makary M A, Daniel M. Medical error—the third leading cause of death in the US BMJ 2016; 353 :i2139 doi:10.1136/bmj.i2139

[12] Bruce, P., Bruce, A., & Gedeck, P. (2020). Practical statistics for data scientists: 50+ Essential concepts using R and python (2nd ed.). O'Reilly Media. p. 141

[13] Kleinbaum, D. G., & Klein, M. (2011). Survival analysis: A self-learning text, third edition (3rd ed.). Springer.

[14]