Exploring Artificial Intelligence Solutions for Tumor Diagnosis with the University Hospital Zürich Otology Department

By
Emilia Casagrande
Alan Healy
Ashwin Pai
Holden Snyder

Exploring Artificial Intelligence Solutions for Tumor Diagnosis with the University Hospital Zürich Otology Department

An Interactive Qualifying Project
submitted to the faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfilment of the requirements for the
degree of Bachelor of Science

by
Emilia Casagrande
Alan Healy
Ashwin Pai
Holden Snyder

Date: 15 October 2020

Report Submitted to:

Dr. Ivo Dobrev
University Hospital Zürich

Professors Ulrike Brisson and Blake Currier
Worcester Polytechnic Institute

Abstract

Otology has few applications of artificial intelligence (AI) compared to other medical fields. We worked with Dr. Dobrev from the University Hospital Zürich (UHZ) Otology department to find where AI could be applied within his team's work. We interviewed otologists at UHZ and AI developers to gather information. We found several areas in the otologists' work where AI could be applied in the far future. We found one area where AI could be implemented in the near future.

Acknowledgements

Our team would like to thank Dr. Ivo Dobrev from the University Hospital Zürich for sponsoring this project. Dr. Dobrev was understanding of many changes in plans and objectives and was a great guide for the project. He was also an excellent resource to answer our questions about otology and provided us with many contacts for our interviews. Without his help this project would not have been possible.

Our team would like to thank the otologists at the University Hospital Zürich (UHZ) for participating in our interviews and answering our questions about otology. Their willingness to remain in contact and answer our questions through emails made this project possible.

We would like to thank those who took extra time to read and suggest revisions to our report, including Anne Healy, Kiran Tremblay, and Dr. Heidi Snyder. Their time was appreciated and increased the overall quality of our report.

Lastly, we would like to thank our sponsors, professors Blake Currier and Ulrike Brisson, for their guidance and support through the project. The professors spent many hours reviewing our report and our findings to improve its quality. Their helpful suggestions both guided and strengthened our report.

Executive Summary

The University Zürich Hospital (UHZ) otology department does not currently incorporate Artificial Intelligence (AI) into their work. Our project aimed to find ways that AI could assist the clinical and research work performed by the otologists. We found that the classification of vestibular schwannoma (VS) tumors is the most applicable area of work within the UHZ otology department for AI.

VS is a non-cancerous tumor that grows on the auditory nerve in the inner ear. Currently, there is a lack of understanding in the causes of VS and the tumor's growth. Many patients diagnosed with VS experience common symptoms of hearing loss and tinnitus. In more severe cases the tumor can press on the brain and pressure critical parts of bodily function. In those cases, surgeons remove the tumor through an invasive procedure. This procedure does not recover the patient's hearing loss. However, this type of surgery is only required if the tumor is actively pressing into the brain. If the tumor is small and the patient is asymptomatic, otologists monitor the tumor instead of operating on it. To monitor the growth of the tumor, an MRI scan is taken at the initial diagnosis and then again six months later. If there is no change in the tumor's size, then the scans are repeated in increasing time intervals.

AI is a classification of computer programs that can modify its own data processing algorithms based on training data. AI can perform many "human" tasks that require decision-making, like image identification and classification. AI requires training data to make decisions. Without the proper training data, the AI cannot make accurate predictions when tested against other data. There are a variety of uses for AI in the medical field. For example, AI has been used to classify cancerous tumors as malignant or benign (Chou et al., 2001, Mangasarian et al., 1995). There are many AI models and systems designed to solve unique problems. For example, an autoencoder is a type of AI that extracts the most important features from an image such that the image could be reconstructed using only those features. Using an autoencoder reduces both the size and number of features within an image, allowing other AI models to easily process the data. In this project, we explore three models that can assist the UHZ otologists in their work when used with an autoencoder. These models are the naive Bayes classifier, the logistic regression model, and the convolutional neural network.

The naive Bayes (NB) Classifier is an AI model that uses Bayes' theorem to predict the likelihood of an event from previous data trends. NB assumes that all features of the data are completely independent of each other (Frank et al, 2002; Xhemali et al, 2009). This assumption makes NB relatively quick and easy to implement (Frank et al, 2002). NB does not require as much training data as other models and is not heavily affected by irrelevant features (Xhemali et al, 2009). NB will largely ignore inputs that do not influence the outcome because of its assumption

of independence. However, the assumption that all features are entirely independent is rarely true, which reduces the accuracy of NB compared to other AI models. Despite this, NB can still be used to observe basic trends in data.

Logistic regression (LR) models use a sigmoid function to predict outcomes based on training data. LR excels at predicting target variables that are classified into specific categories with a given confidence. Previously, LR has been used in the medical field successfully to predict whether breast cancer tumors are malignant or benign (Mangasarian et al, 1995). Similar to NB models, LR is relatively simple to construct and implement. LR struggles when there is a wide range of output categories. LR should not have difficulty in classifying tumors for the otologists at UHZ as there are only 2 categories, likely to grow and not likely to grow. The model gives a confidence rating when categorizing input. When this rating is converted to a percentage, it can be used by doctors to interpret the likelihood that a tumor will grow. LR is slower to set up than NB, but it produces more accurate results (Frank et al, 2009). It also requires more data preparation compared to similar models.

Convolutional neural networks (CNN) are a model specialized for pattern recognition in images. A typical CNN consists of an input layer, output layer, and many hidden layers. The input layer contains the image to be analyzed. The output of a CNN is typically a list of probabilities that the image contains certain features. A CNN's hidden layer contains several preparatory layers. As the image is processed by these layers, matrices of weights known as filters are applied to identify features in the image, in a process called convolution. Then, a process called pooling is applied to remove unimportant features from the image. The process of convolving and pooling is repeated multiple times, with increasingly sophisticated filters to extract more complex features. After multiple iterations, the extracted features are passed to a classification layer. The classification layer processes the extracted features together to classify the image. For example, if the features are a car tire, car headlights, and a car door, the classification layer would classify the image as a car. CNNs typically have more accuracy than LR and NB models. However, CNNs also require a larger set of training data than LR or NB models to achieve high accuracy. The increased complexity of CNNs may necessitate more time to implement and train.

To complete this project, we interviewed several otologists from UHZ to understand their work. From information gathered in these interviews, we identified possible areas where AI could be implemented at UHZ's otology department. After further research, we concluded that the classification of VS tumors was the most applicable area for AI. The large amount of training data available may allow for an effective AI to be trained. In other medical fields, AI models have already been created for tumor classification; we hypothesize that such a model could be used to classify VS tumors. Our team then conducted follow-up interviews with an otologist who operates on VS tumors to gather more information on his work. This information helped guide our research on the types of AI that could best assist in tumor classification. This research included conducting interviews with AI experts in Switzerland and the United States. To communicate this information, we developed a guidebook for otologists at UHZ on how to approach AI development within VS diagnosis. This guidebook informs the otologists on the basics of AI, explains the types of AI best suited to tumor classification, and lists some companies from the local Zürich area that specialize in AI technologies.

Recommendations

We recommend using patient medical history and MRI image data to train an AI to classify vestibular schwannoma tumor growth. The factors contributing to VS tumor growth are not yet fully understood. Aspects of the patient's medical history could be contributing factors to the growth of the tumor. Care must be taken to ensure that this data remains anonymous, which can add additional time to the data cleaning stage of development. We think the benefits of including this data outweighs the costs. The data contained in patients' medical history could allow otologists to identify additional risk factors that contribute to tumor growth. Using this data to train the AI could allow for a more accurate model.

We recommend ensembling multiple different AI models with an autoencoder to identify tumor growth. The autoencoder reduces the number of features in the MRI image to only the most important ones. Three-dimensional MRI images may be too complex to process with multiple AI models without some form of compression. An autoencoder improves the performance of successive models, especially logistic regression (LR). Ensembling is a process of using multiple AI models in parallel and combining their results. By combining the results of the models - or selecting the most accurate model - otologists can achieve more accurate results.

We recommend the otologists at UHZ use a naive Bayes (NB) Classifier as the first model created to identify tumor growth. Choosing this model first would allow the otologists to check the feasibility of the project without investing as much time. The advantages of NB are that it is fast, easy to implement, ignores irrelevant features, and does not require as much training data as other models. However, NB makes significant assumptions about the data that lead to accuracy issues. If NB can predict tumor growth, then the otologists will know that the prediction is possible. Then, more accurate models can follow NB to improve upon the predictions. NB is best used as a test model to ensure that it is possible for VS tumor growth to be predicted using AI.

We recommend the otologists at UHZ use a logistic regression (LR) model if the naive Bayes model can classify tumor growth but is not accurate enough to be used clinically. LR is much faster to create and implement compared to convolutional neural networks (CNN), but is slower than NB. LR has been used to predict tumor malignancy in the past (Chou et al., 2001, Mangasarian et al., 1995). The advantages of LR include quick setup, effective results, and the ability to train the model as it is used. The disadvantages of LR include its low accuracy compared to CNN, and that it is prone to overfitting if the data is not varied. LR serves as an intermediate model between CNN and NB. If LR is accurate enough to be used clinically, there is little reason to advance further to CNN.

We recommend the otologists at UHZ use a convolutional neural network if the logistic regression model is not accurate enough to be used clinically. A CNN is an AI model that takes longer than others to implement but is very powerful. CNNs have especially strong pattern recognition, which can make them accurate enough for clinical use. CNNs have been used in medical research before, with successful results. The disadvantages of CNNs include that they are time-intensive and costly to implement, that they require more training data than other AI's, and that they are prone to overfitting if the data is not varied enough. While CNNs are very powerful, they may be more than the otologists need. This model should be used if LR fails to predict tumor growth or is not accurate enough to be used clinically.

Contributions

Alan Healy. Alan was the primary author for background information on artificial intelligence and machine learning concepts. Alan was the primary researcher for convolutional neural networks, and primary author for related sections. Additionally, Alan led the interview with an AI expert. During other interviews and weekly meetings, Alan took notes and summarized key information learned. Alan also managed the timing of team presentations. Throughout the writing process, Alan edited other sections of the report for spelling, grammar, and content. Alan designed the format of the deliverable guidebook for otologists. Finally, Alan designed the overall formatting of the report.

Ashwin Pai. Ashwin served as a leader for several interviews with the otologist and the primary interviewer for the AI companies accompanied by both Alan and Holden. In addition to this, Ashwin also was the primary author for the sections describing the field of otology and the common ailments experienced in the ear and objective two of the methods chapter alongside Emilia. For other sections, Ashwin served as the primary editor, going through and fixing problems in tone, flow and sentence structure. These first edits were then followed up and modified by all other members of the team. Lastly, Ashwin worked on researching and describing the companies found in the guidebook.

Emilia Casagrande. Emilia served as the primary note taker for several interviews with both AI companies and otologists and all meetings with our sponsor. Emilia was also responsible for researching and scheduling interviews with AI companies in Switzerland. In order to ensure a quiet place to meet, Emilia booked rooms on campus each day. In addition, Emilia assisted in writing the Background and Methods sections with Ashwin. Emilia was also responsible for editing multiple sections of the report including the Introduction, Findings, and Conclusions and Recommendations. These edits focused on grammar, structure, flow, and punctuation. Lastly, Emilia also served as one of the primary contributors to the glossary of terms.

Holden Snyder. Holden served as a lead interviewer for several interviews with otologists with Ashwin. Holden was co-author of the introduction and the deliverable guidebook with Ashwin. Holden wrote the artificial intelligence section of the guidebook and was responsible for finding tutorials for the AI models. He was the primary researcher for the naive Bayes and logistic regression AI models, and was the primary author of all sections pertaining to those models. Holden was a primary author of the findings and conclusions sections. Holden served as a secondary author of the background section. Holden was an editor for the introduction and methodology sections of the report, as well as a citation checker.

Table of Contents

List of Figures

List of Tables

Glossary of Terms

Artificial Intelligence (AI)	A class of computer programs that can modify their own algorithms
Artificial Neural Network	the underlying architecture for many AI systems
Auditory Nerve	A nerve in the inner ear that connects the cochlea to the brain
Auto-encoder	A type of AI that reduces images to their most important features
Back-Propagation	A process used to adjust the AI's weights and minimize error
Benign	A tumor that will not continue to grow

Cochlea	Spiral shaped organ in the inner ear
Computed Tomography (CT)	Imaging technique that uses X-rays to look at parts of the body
Conductive Hearing Loss	Hearing loss due to damage to tympanic membrane
Convolutional Neural Network	Type of artificial neural network specialized for image recognition
Cystic	Tumors containing fluid
Data Cleaning	The process of preparing data before it is fed to an AI model for training.
Edges	Connections between nodes where data is passed from one node to another with a weighted value
Electrocochleography	Recording electric potentials generated in the cochlea in response to stimulation
Ensembling	Using multiple AI models either in succession or in parallel
Hidden Layer	Layers of data processing that take place between the input and output layers of the AI model
Logistic Regression (LR) Classifier	A model of AI that uses a Logit function to classify inputs
Machine Learning	The application of AI to a problem
Magnetic Resonance Imaging (MRI)	Imaging technique that uses magnetic fields to look at parts of the body
Malignant	A tumor that will continue to grow
Naive Bayes Classifier	A model of AI that uses a modified Bayes' Theorem to classify inputs
Nodes	Structures that hold data and perform some mathematical operation on the data
Otology	The study of ears and their function
Overfitting	Inaccuracy introduced to an AI model when training data does not have enough variance
Sensorial Hearing Loss	Hearing loss due to damaged nerve endings in the cochlea

Training Data	Data used to train an AI
Tumor	Abnormal cell growth
Vestibular Schwannoma	Non-cancerous tumor that grows in the inner ear close to the brain

Chapter 1: Introduction

Artificial Intelligence (AI) has seen novel uses in many medical fields; it is most prominently used in diagnostics and imaging (He, 2019; Liu X. 2019). AI has been used in diagnostic applications, data analytics, and robotic surgeries. However, not all medical fields have benefited from the integration of AI. The study of ears and their function, known as otology, has a noticeable absence of AI technologies compared to other medical fields (Chung, 2017). This problem does not stem from an aversion, but rather an unawareness of AI and its capabilities. Our sponsors, the otologists at University Hospital Zürich (UHZ), are interested in finding the ways AI can be implemented into their work. This report discusses how AI can be applied in the classification of vestibular schwannoma tumors.

Vestibular schwannoma (VS) is a type of tumor that grows in the inner ear close to the brain. While the tumor is straightforward to identify using magnetic resonance imaging (MRI), there is currently no technique that can predict if the tumor is likely to grow. Treatment of the tumor varies based on the size and symptoms experienced by the patient. If the tumor is large and causing symptoms, it is either surgically removed or treated with radiation. Alternatively, if the tumor is small, it is monitored by taking scans every 6 months to see if any properties of the tumor have changed (C. Röösli, personal communication, September 28, 2020). Our project found that AI could be a powerful tool to classify VS tumors as likely to grow or not. A classification tool would increase the efficiency of the diagnostic process, since patients would not have to wait six months or longer to know if the tumor is growing. This would allow doctors to remove growing tumors before there is more damage to the ears and brain.

It is important to understand that our project did not focus on creating an AI for the field of otology. Instead, we worked with UHZ researchers to understand their work and find areas where AI could assist. Currently, there are very few AI solutions that are tailored specifically to otology. The goal of our project is to find the areas of the researcher's work that could benefit from AI and inform our sponsors on the most applicable types of AI. Currently, there are a variety of AI models that specialize in tumor classification. For example, one study tested the accuracy of AI in diagnosing breast cancer tumors as either malignant or benign. The study found that the AI system accurately identified 91.0% of malignant tumors (Chou et al. 2001). While there are many AI models created to diagnose different tumors, currently we are unaware of any that assist in otological tumor classification.

AI is not as widely used in otology as other medical fields. Tumor classification is a common research field in various regions of the body. However, not all medical fields benefited equally from the integration of AI. Tumors found in the ear are not as common and therefore have not required the construction of an AI model. Out of every million people, 19 to 23 will be diagnosed with a VS tumor in the inner ear (Stangerup, 2012). While this tumor is quite rare, there are severe complications if it is untreated such as hearing loss and tinnitus. Identifying a tumor's presence is quick, however, classifying it as likely to grow or not is a lengthy process. Patients diagnosed with VS tumors must return for MRI scans after 6 months to monitor the growth. A technology that could predict if a tumor would grow could increase the efficiency of the treatment process for doctors and decrease the number of MRI scans needed to monitor the tumor's growth.

The goal of our project is to discover the types of AI best-suited to assist the UHZ otology department and inform the researchers on these technologies. To accomplish this, we interviewed researchers from the UHZ otology department to understand their work. Additionally, to gain an understanding of suitable AI technologies for the otologists' areas of research, our team interviewed both local and international AI companies and developers. The results of this report will be used to inform the UHZ otology department on the feasibility of AI systems into their clinical and research work. Our project focused specifically on the integration of AI technologies in the classification of VS tumors. The contents of this report will help guide the future otologists in the treatment process of VS tumors, potentially leading to better patient outcomes.

Chapter 2: Background

In this chapter, we begin with a discussion on the field of artificial intelligence (AI) and its current applications within the medical field. Next, we explore otology, including the physiology of the ear, common diagnostic techniques, and data analysis methodologies. Finally, we introduce our sponsor, Dr. Ivo Dobrev and the University of Zürich Hospital Otology Department, along with the problems they are attempting to better understand.

2.1 Otology

In this section, we introduce the field of otology. Here, we discuss the anatomy of the ear along with each structure's function. We describe the tools used to identify the presence of otological ailments. Otology is the study of ears and their function. This includes accompanied diseases and disorders along with their diagnoses and treatments. The field itself is diverse, with research topics ranging from improving cochlear implants to new otologic surgery techniques and technologies.

2.1.1 How does an ear function?

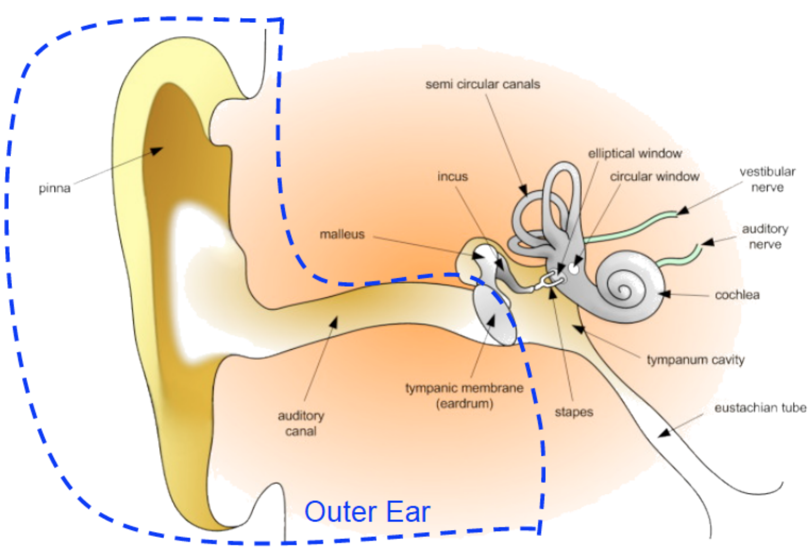


Figure 1: Diagram highlighting the outer ear (McCarthy).

The outer, middle, and inner ear are the three primary sections of the ear. Each region plays an integral role in sound processing. The outer ear, which includes the pinna, ear canal, and eardrum, is responsible for detecting and directing vibrations. Figure 1 shows a diagram of the ear with the outer ear highlighted. The pinna, the portion of the ear that protrudes from the skull, funnels sound waves from the environment into the auditory canal. As sound travels through the narrow passage, the vibrations are continually amplified as the sound waves reflect off the walls of the ear canal. At the end of the ear canal, the vibrations meet the tympanic membrane, commonly known as the eardrum. This thin layer of connective tissue is crucial to the intricate relay that transmits airborne sound to our fluid-filled inner ear (Mozaffari et al., 2019). Noise hits the eardrum causing it to resonate, triggering the oscillation of the ossicular bones in the middle ear.

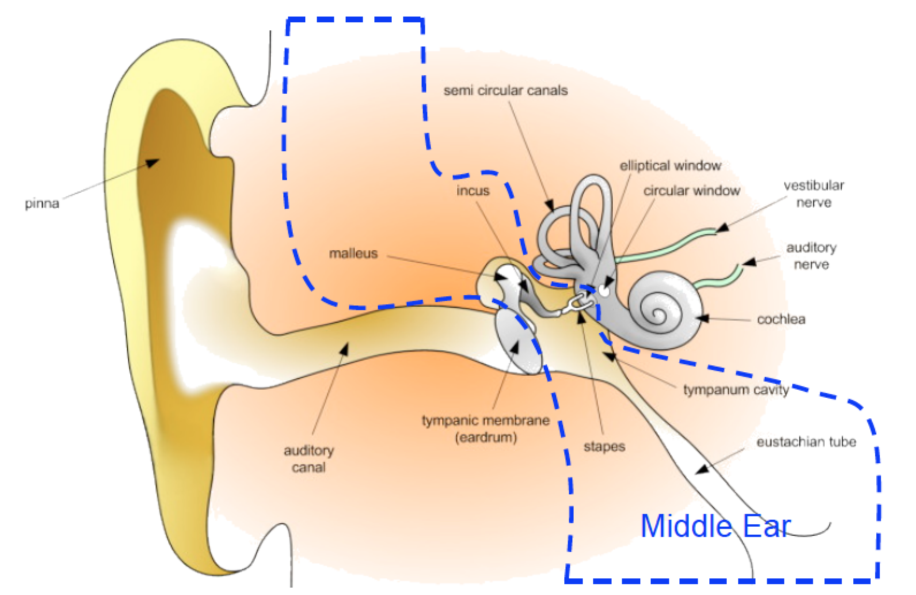


Figure 2: Diagram highlighting the middle ear (McCarthy).

The middle ear, which is composed of the ossicles and oval window, is responsible for the translation of sound into pressure waves. The parts of the middle ear are highlighted in figure 2. As sound penetrates the eardrum, it causes the movement of the ossicular bones. The ossicular region comprises the three smallest bones in the body; the incus, stapes and the malleus. The hammer-like structure, called the malleus, is the first bone in the ossicular region. This bone is connected to the eardrum. As sound vibrates the tympanic membrane, the resulting force causes the stimulation of the malleus. As the malleus moves back and forth like a hammer, it transmits its energy to the incus, which then passes it to the stapes. When the three bones are moving in unison, the ossicular region moves like an engine piston (Stomackin, et al. 2019). The stapes, the end of the piston, oscillates stimulating the oval window. This membrane acts as the protective layer to the fluid-filled cochlea by translating vibrations into pressure waves.

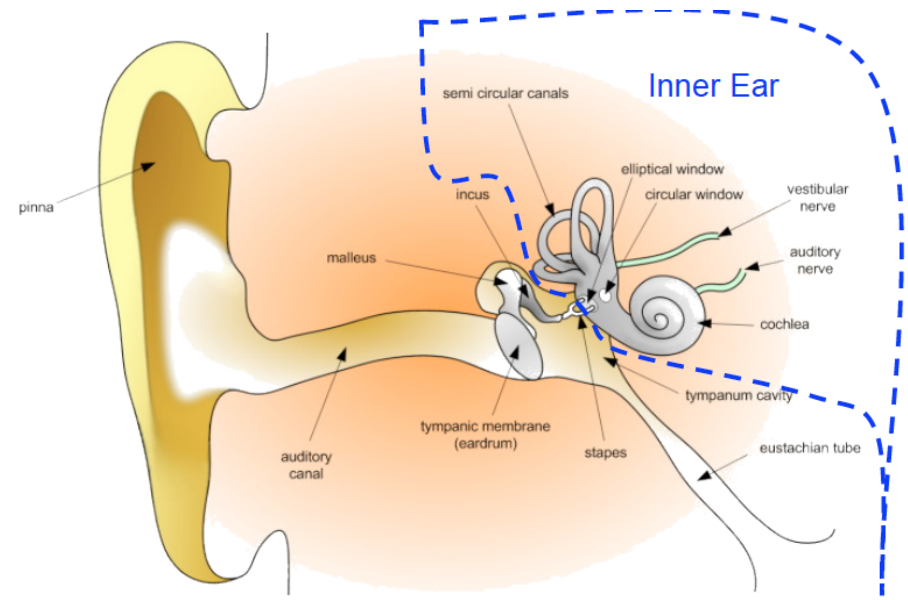


Figure 3: Diagram highlighting the inner ear (McCarthy).

The final stage of sound processing takes place in the inner ear, as shown in Figure 3. This stage involves the translation of pressure waves into electrical impulses using the cochlea and the auditory nerve. As pressure waves are generated by the oval window, they enter the snail-shaped structure known as the cochlea. The pressure waves set the endolymphatic fluid of the cochlea into motion, similar to ripples when a stone is dropped in a puddle. As the fluid vibration travels through the cochlea, it interacts with a hearing receptor called the spiral organ. The spiral organ translates pressure waves into electrical signals (Lawrence, 1966). The impulses created by the cochlea are sent into the brain via the auditory nerve. The brain associates patterns of electrical signals with recognizable sounds, such as a car engine or dog bark.

2.1.2 Ailments within the Ear

According to the Global Burden of Disease Study, a comprehensive global study of ailments and their effects, hearing loss is the fourth most common disability in the world. In fact, in the United States, hearing impairment affects 50% of people between the ages of 60 to 69 and 80% of people 85 and older (Cunningham, 2017). As life expectancy continues to increase around the globe, the number of hearing impairment cases will increase (Cunningham, 2017). Currently, the two most common types of hearing loss include conductive, and sensorial (Zahnert, 2011). The following section discusses these ailments and their respective causes.

Sensory loss is the most common form of hearing impairment and its severity can be directly attributed to the health of the cochlea (White et al., 2020). The cochlea of a patient suffering from sensory hearing loss contains missing or damaged nerve endings in the spiral organ. When these nerve endings are not functioning properly, incoming sounds cannot be converted into electrical impulses (Zahnert, 2011). This prevents the patient from hearing or recognizing sounds. The most common risk factor of sensory hearing loss is old age. According to the Official Journal of the Italian Society of Otorhinolaryngology, 18% of women and 35% of men between the ages of 60-69 suffer from sensory hearing deprivation (Ralli et al., 2017). However, when the range of surveyed ages was increased by just 15 years, researchers saw that 71% of women and 84% of men between the ages of 75-84 saw decreased hearing (Ralli et al., 2017). As a person ages, the nerve endings within their ear begin to deteriorate and eventually stop working. This is a process that occurs through normal deterioration as the functions of the body begin to decline. Since the nerves in cochlea are non-regenerative, as a person gets older the volume and clarity at which they hear declines (White et al., 2020).

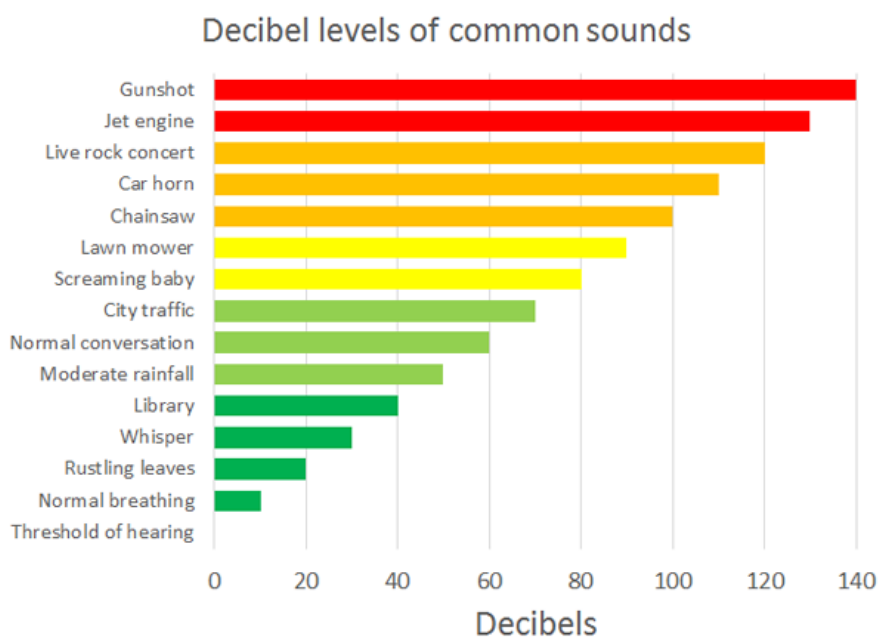


Figure 4: Decibel levels of noises that people may encounter (Mehta et al., 2012)

Conductive hearing loss is another type of hearing loss caused by damage to the tympanic membrane. Conductive hearing loss occurs when the tympanic membrane is damaged or completely ruptured (Liberman et al., 2015). The cause for this type of damage is often determined by the patient's age. Usually when the patient is older, this is the result of either frequent or sudden exposure to loud noises. Ears are built to handle noises ranging from 0 to 80 dB (National Institute of Deafness and other Communication Disorders, 2014). Loud sounds are dangerous because of the strain they put on the ear canal and drum. Sound levels above 80 dB carry enough force to break or loosen the tympanic membrane (National Institute of Deafness and other Communication Disorders, 2014). When damage occurs to the eardrum, sound cannot be properly conducted through the ossicular region. As such the cochlea does not create these electrical impulses.

Tumors are another disease found in the ear. A tumor is a general word for mass caused by a neoplasm, or more simply put, the abnormal growth of a cell (Baker et al., 1965). If a tumor is cancerous, that indicates that a mass has formed and the cells are continuing to divide in an unregulated manner. These types of tumors are classified as malignant and they must be treated in a timely fashion. If the mass is left untreated, there is an increased chance of the tumor metastasizing and spreading to other regions of the body (Baker et al., 1965). In comparison, if a tumor is benign, or non-cancerous, this is an indication that the mass is not increasing in size and is primarily localized to a specific region of the body (Steeg, 2006). Benign tumors are not as dangerous as malignant tumors, however, if the mass remains undiagnosed, it can lead to other health complications such as tumor growth within the middle and inner regions of the ear.

Vestibular schwannoma (VS) is a non-cancerous tumor that grows on the auditory nerve in the inner ear. Nerves are similar to wires in that they are required to be insulated. Nerves are wrapped in a coating of Schwann cells which protect the electrical impulse. VS is a disease that causes this coating of cells to rapidly divide and form a mass on the auditory nerve (Greene & Al-Dhahir, 2020). The weight of this mass can cause the auditory nerve to function improperly leading to hearing loss, tinnitus, and balance loss. Since this type of tumor is relatively slow growing, it often goes for long periods of time without being diagnosed. For example, this type of tumor affects a patient's hearing ability. Since this disease is commonly found within the same age group as hearing loss (ages 50 - 60), doctors often misdiagnose the tumor as a standard case of hearing loss (Greene & Al-Dhahir, 2020). Only with further imaging techniques, such as Magnetic Resonance Imaging (MRI), can the tumor be identified and surgically removed.

Temporal bone cancer is another disease that affects the middle and inner ear. The temporal bone is located on the side of the skull, almost exactly where the ears are located. It is common to see that the cancer has metastasized to the ears before formal diagnosis has occurred (National Institute of Deafness and other Communication Disorders, 2017). In a study conducted by medical researchers at Oxford University, 17 people were seen to have tumor growth located in the ear. When tested for temporal bone cancer, 13 of the subjects were also seen to have tumor growth on the skull (Martinez-Devesa et al., 2008). This type of tumor is cancerous and therefore must be treated quickly. When the tumor is present in the middle ear, the patient can experience sharp pain in the ear canal, perforation of the eardrum, and the dissipation of fluids through the canal. The diagnosis of this cancer is straightforward. Similar to a VS tumor, a doctor will request an MRI scan. However, once the MRI confirms that tumor growth is present, the treatment varies. Temporal tumors that have spread to other sectors of the body indicate that the tumor is aggressive. Some doctors argue that radiation therapy is the best approach to take when removing the tumor as it is the most effective at treating large spreads of cancerous cells (National Institute of Deafness and other Communication Disorders, 2017). However, exposure to this type of treatment can affect the nerves within the ear resulting in the degradation or complete loss of function of the middle and inner ear. The other alternative to this type of therapy is the traditional surgical removal of the tumor. The issue with this approach is that it requires a surgeon to operate within the already sensitive ear region (Martinez-Devesa et al., 2008). In order to decide what approach mitigates the spread of the cancer and protects the function of existing structures, advanced imaging is required.

2.1.3 Diagnostics Within Otology

A common method of obtaining the medical data needed to diagnose ailments is medical imaging. Methods of imaging within otology include Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI). CT scans are three-dimensional x-ray images of the body. The CT machine aims a beam of x-rays at the patient. The beam is then quickly rotated around the patient. An x-ray detector is positioned opposite from the x-ray emitter while it rotates. The x-rays that leave the patient on the other side are picked up by the detector (van Beek, 2008). The machine then plots the x-rays that made it through on a grayscale to make a two dimensional "slice" of that section of the body. These slices can be viewed individually or fused together to create a three-dimensional image of the body.

MRI is an imaging method that can be tuned to detect different types of tissue or bone (Berger, 2002). Compared to a CT scan, an MRI allows for more distinction between tissue types (van Beek, 2008). An MRI creates a three-dimensional image very similarly to the CT scan, by taking slices of the body and combining these slices into a three-dimensional image. The scan itself uses a powerful magnet to create a field that aligns the spins of the hydrogen atoms within the body (Berger, 2002). Radio waves that resonate with hydrogen are then sent into the patient to excite the atoms. The radio wave emitter is turned off and the hydrogen atoms are allowed to gradually return to a normal state. When the atom returns to a normal state, it emits a photon. By tracking where these photons came from, their intensities, and the time it took for them to be emitted, the machine can form an image of the body part in question (Berger, 2002; van Beek, 2008). Different body parts have different emission timings when the radio wave emitter is turned off. By suppressing all emissions that come outside of a certain time period, the scan can be tuned to look at specific parts of the body. For instance, by only tracking emissions known to come outside the emission window for fat, the machine can take an image without scanning for fat (Berger, 2002).

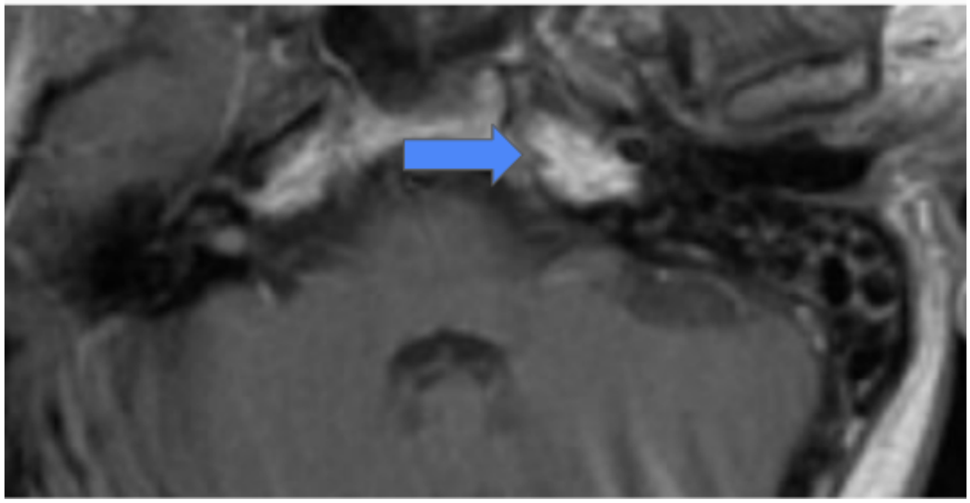


Figure 5: Small vestibular schwannoma tumor, .5cm diameter (Hain, 2020)

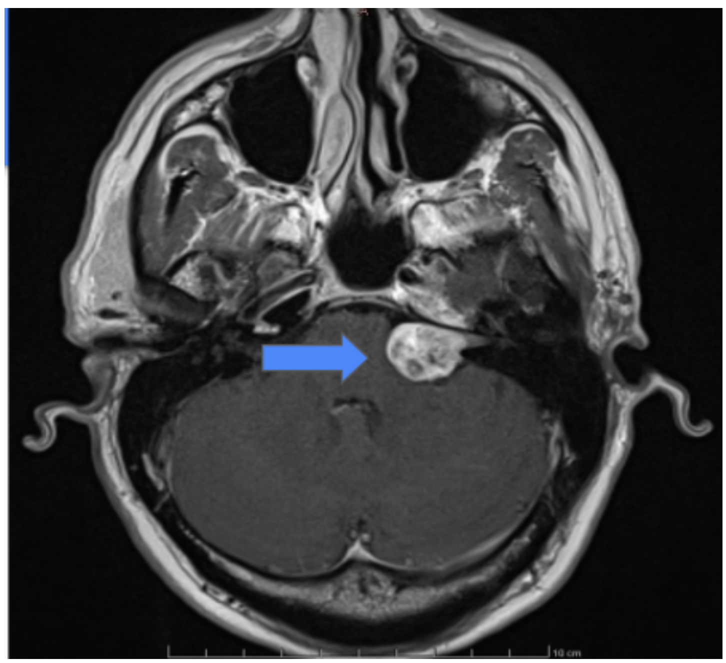


Figure 6: Large vestibular schwannoma tumor ~3cm diameter (Hain, 2020)

CT and MRI scans are the most effective tools in identifying a tumor's presence. However, they are seldom the first form of diagnosis because of their expensive cost and limited availability. Often, doctors will request the use of medical imaging only when the symptoms of the patient have indicated that tumor growth may be present. Typically, doctors will perform a physical test on the patient, looking for lumps, changes in skin color, or any other skin abnormalities (Hamilton, 2010). If there are no physical signs of damage, doctors will perform serologic analysis of the patient. The doctor will take a sample of blood from the region in question, and then mix the blood with a tumor marker agent. If a tumor cells are present in the blood, it will be visible once the sample is analyzed.

Once a tumor has been detected, doctors must determine whether the mass is likely to grow or not. Currently, there are no tools that can diagnose whether a tumor is likely to grow or not from just one image. Instead doctors take multiple scans over a set amount of time to see if the tumor has increased in diameter (Hain, 2020). By comparing the two images over a set amount of time, doctors can learn the rate of the tumor's growth, along with its probability of metastasis. Figure 5 and figure 6 show sample cases of VS tumors from different patients that vary in size.

Hearing Threshold Increases for Varying Degrees of Hearing Loss

Degree of Hearing Loss	**Auditory Threshold Increase (dB)**
Mild	25 to 40 dB
Moderate	40 to 55 dB
Severe	70 to 90 dB
Profound	90 dB and beyond

Table 1: Auditory threshold chart showcasing the varying thresholds that determine hearing sensitivity.

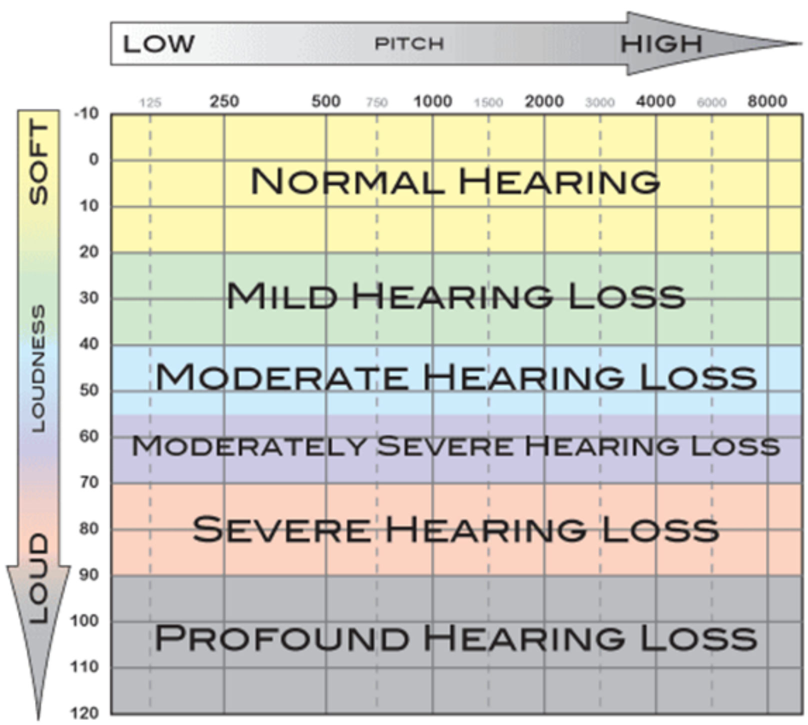


Figure 7: Sample audiogram detailing the ranges and names of hearing loss degrees (Balasubramanian, 2015).

In order to understand hearing loss diagnosis, it is important to discuss the measurement of hearing sensitivity. To determine the strength of an individual's ear the auditory threshold is measured. Briefly speaking, the auditory threshold is the lowest volume at which the ear can detect vibrations; for average humans it is around 0 dB, otherwise known as near-silence (Westman & Walter, 1981). However, when a person is experiencing hearing loss, this threshold shifts and increases. The number of decibels the threshold shifts by is directly correlated to the severity and degree of the individual's hearing loss (Holman & Drummond et al., 2019). To measure the auditory threshold of a person, doctors will administer an air conduction test. When ears detect a sound, they return a signal back into the environment. Air conduction testing monitors the return signal from the ear to determine whether a sound was registered (Holman & Drummond et al., 2019). In air conduction testing, a wide variety of frequencies that are used in human communication are propagated to the ear through headphones to determine whether the ears can register the vibrations as noise. The results from the air conduction testing indicate the softest volumes and frequencies at which a person could register sound. This data is put into a graphical form which is known as an audiogram. Based on the results of the graph, the change of the auditory threshold can be measured as well as the severity of hearing loss. There are 4 degrees of hearing loss: slight, mild, severe and profound (Bance, 2007). Table 1 shows the results of an auditory test conducted on a person with severe hearing loss, and another person with no hearing loss. There is a shift in the auditory threshold of around an 80 dB increase in intensity for the patient who is suffering with severe hearing loss. According to the National Hearing Test Organization, the 4 hearing loss types are produced by the changes in auditory threshold shown in figure 7.

2.2 What Is Artificial Intelligence?

The field of artificial intelligence (AI) involves the creation of computer programs that solve problems thought to require human intelligence. Many tasks performed by humans, such as language processing and image recognition, are said to require some degree of intelligence due to their complex nature (Nilsson, 2014, 1). Over time, computer scientists and mathematicians have improved AI techniques to perform these tasks (Nilsson, 2014, 2). Currently, many fields employ AI to handle complex tasks, including finance (Bahrammirzaee, 2010, 1165-1166) and medicine (Rajkomar, M.D. et al., 2019, 1347). In order to consider how AI could be applied to otology, we must first discuss how AI differs from "traditional" algorithms, as well as how programmers can solve problems with AI.

In Computer Science, a traditional algorithm is a series of instructions that a computer performs to transform input data into a result (Ismail, 2018). Each instruction of an algorithm must be possible to perform with a computer and unambiguously process a finite amount of input data

(*Higher Mathematical Jargon*, 2019). Programmers must explicitly define instructions, through conditional statements and mathematical operations. The following is an example of an algorithm, drawn from an example by A. M. Kuchling.

Goal: Given a finite list of positive integers, *L*, find the maximum number in the list.

Example Algorithm:

1. Create a variable to hold the current maximum value, *M*.
2. Set *M* to 0.
3. For each number, *N,* in the list *L*, do the following:
    a. Compare *N* to *M.*
    b. If *N* is greater than *M*, then set *M* to the value of *N*.
4. *M* is now the maximum value of the list. Return *M* as the result.

The above algorithm is easy to implement in a computer program. It is guaranteed to terminate, as the provided list *L* is finite, and computes one comparison for each element in the finite list (Kuchling, 2012). Furthermore, each step of the algorithm is explicit and unambiguous. In conventional programming, algorithms can be combined and altered to reach high levels of complexity. However, each step of the underlying structure is still explicitly defined by programmers.

Conversely, AI algorithms typically have some elements that are not explicitly defined by programmers. AI systems can modify their algorithms based on sample data (Rajkomar, M.D. et al., 2019, 1348). As developers pass example inputs to an AI system, they can evaluate the system's performance. If unsuccessful, the AI system will modify its algorithms to achieve a better result. This process is known as training. With sufficient training, AI systems can become highly accurate in areas such as classification, identification, or prediction (Rajkomar, M.D. et al., 2019, 1348).

Both AI and conventional algorithms have advantages and disadvantages that affect their use cases. AI can streamline complicated processes and analyze multidimensional data quickly. It can also identify complex data patterns previously unrecognizable by humans (Hosny et al., 2018). However, AI requires a large collection of training data to achieve a high level of accuracy. AI models tend to amplify biases present in the data, which may cause higher rates of failure when use-cases differ from the training data (Rajkomar, M.D. et al., 2019, 1355). Furthermore, the self-updating nature of AI systems reduces their "explainability"; it can be difficult to describe *how* an AI calculates a result, especially if it operates on many variables (Hosny et al., 2018). Conventional algorithms, however, allow for more transparency and explainability in the code. By following the algorithm step by step, data scientists can identify and fix problems quickly, rather than having to retrain a model with different parameters. Still, conventional algorithms require all logic to be explicitly defined by programmers, which means the programmers must plan out exactly how the data will be processed beforehand. When problems are novel and not yet fully understood this may not be possible.

2.2.1 Machine Learning

Machine learning (ML) is the process of applying AI techniques to solve problems (Rajkomar, M.D. et al., 2019, 1347-1348). Some ML structures are more suited to classification of data, while others are useful for predicting trends. For example, convolutional neural networks (CNN) add extra layers of data computation that are well-suited to image processing. CNNs allow for the identification and separation of features (Brinker et al., 2018). A feature refers to any pattern in data, such as shapes, objects, or trends. When approaching a problem with the goal of using ML, the following process can be helpful.

- **Define the problem domain.** In order to work with an AI system, it is important to thoroughly understand the problem domain. Factors such as data availability, variety, and complexity will inform the choice of which ML models are feasible to use.
- **Gather and clean the data.** For the system to work properly, data must meet certain criteria. All data used to train the machine must be in the same medium and in a consistent format. It may even be necessary to apply some preprocessing to the data, such as removing empty space around images or blurring high resolution images. This process is known as data cleaning.
- **Train and test the model.** After creating a ML model and training it on sample data, AI can uncover patterns in the data that were not previously visible.
- **Analyze the discovered patterns.** Developers can analyze and interpret the patterns identified by ML systems to come to conclusions.
- **Communicate results.** Developers can communicate the results and possible applications of the patterns found through machine learning.

By using such strategies, complex machine learning problems can be broken down into more manageable parts, allowing for the development of effective machine learning models.

Creating an accurate AI system requires careful training from varied data. The process of training involves multiple steps. The first step in training is to acquire data which has already been classified. This may require manual classification of a data set, but once completed, the AI can begin training. The creators of the AI can run it on some portion of the training data. After each input is processed, a "loss function" determines to what extent the algorithm's output was accurate.

Based on the output of this loss function, a process called backpropagation is used to adjust the AI's weights and minimize error (Bahrammirzaee, 2010, 1167). A weight is the importance the machine has assigned to a specific feature. If the machine finds strong correlation between a specific input and output, the weight for that input will increase for that output. The specific details of the backpropagation algorithm vary for each AI. As this sequence of processing data and adjusting weights is repeated, the AI becomes more accurate. More training data usually means more accurate outputs from the AI. However, too much similar training data can lead to accuracy issues.

Training an AI from data that is too similar causes a problem known as overfitting. Overfitting is inaccuracy introduced to the model when trying to classify or predict anything not closely related to the training data. For example, if a developer designed an AI to recognize pictures of dogs, then trained it exclusively on dogs with a solid coat, it would fail to recognize dogs with a spotted coat. The importance of a solid coat was emphasized to the AI, since it was a consistent feature of the training data. With more diverse images of dogs, such overfitting would be less likely to occur. Therefore, diverse training data is necessary to implement accurate machine learning models.

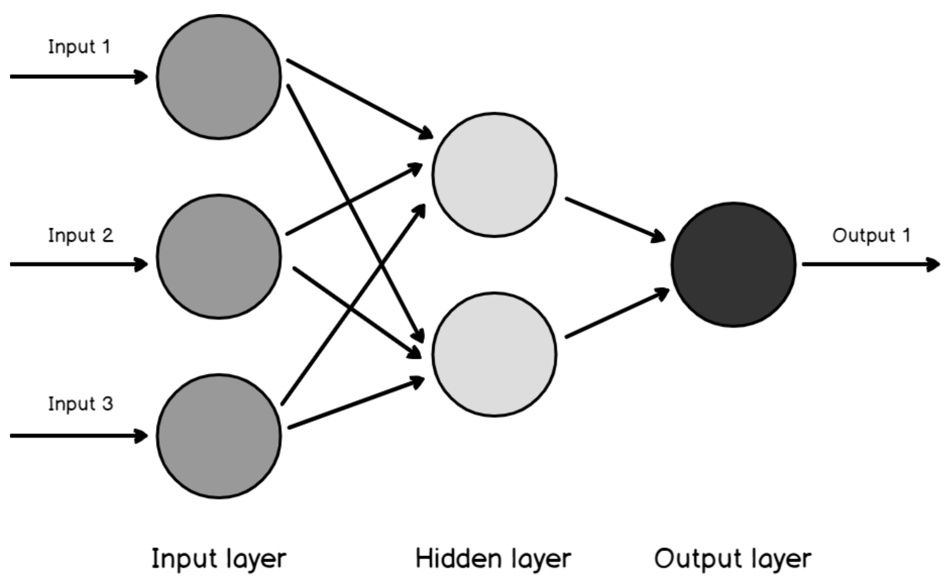


Figure 8: Visual representation of a basic artificial neural network (Asanka, 2020).

Artificial Neural Networks (ANNs) are the underlying architecture for many AI systems (Bahrammirzaee, 2010, 1165-1166). ANNs consist of a collection of nodes and edges. Nodes are structures which hold data and perform some mathematical operation on that data. Edges are connections between nodes, where data is passed from one node to another with a weighted value (Brinker, et al., 2018). In figure 8, the circular structures represent nodes, and the arrows between circles represent edges. Figure 8 also demonstrates the organization of nodes: they exist in a series of connected layers that form a network. Generally, a neural network will contain an input layer, where each node corresponds to a variable that is passed into the network for processing (Brinker, et al., 2018). The input layer passes data into one of many hidden layers, where most of the data processing occurs (Brinker, et al., 2018). Finally, the last hidden layer passes data to an output layer, which contains the result(s) of the calculation.

The accuracy of a neural network is largely determined by the weights of edges. Broadly, weights are values within a range (such as 0 to 1) which determine which features are important in the final calculation of a neural network (Bahrammirzaee, 2010, 1165-1167). For example, if a neural network is designed to differentiate between different breeds of dog, features such as color and height may be deemed important, and therefore would have higher weights. A single neural network can have thousands of different weights, which precludes programmers from manually changing these weights. Rather, the network is trained on a set of input data with known output values. Accurate AI models will have weights of edges that quantify the importance of a feature to a data set.

Naive Bayesian (NB) models are ML models based on using Bayes' Theorem to classify data. Bayes' Theorem is

$$p(y_i|x_1, x_2 ... x_n) = \frac{p(x_1, x_2 ... x_n|y_i) * p(y_i)}{p(x_1, x_2 ... x_n)} \quad (1)$$

where $x_1, x_2 ... x_n$ are the features of the data, and y_i is a potential classification. $p(y_i|x_1, x_2 ... x_n)$ represents the probability that the features of the data represent something from class y_i, $p(x_1, x_2 ... x_n|y_i)$ represents the probability of these features occurring in the data if the classification is assumed to be true. $p(y_i)$ is the independent observed probability of class y_i occurring, and $p(x_1, x_2 ... x_n)$ is the probability of the features of the data occurring together. The naive Bayes model assumes with Bayes' Theorem that all features from the data are completely independent of each other (Yildirim, 2020). This modifies Bayes' Theorem to look like this

$$p(y_i|x_1, x_2 ... x_n) = p(y_i|x_1) * p(y_i|x_2) * ... p(y_i|x_n) \quad (2)$$

In this case, the probability of a given feature is calculated separately from all other features. A real-world example of this difference would be like the probability of pulling a specific card from a deck of cards. The initial probability of pulling a specific card is $\frac{1}{52}$, and the probability of pulling that same card on the next pull becomes $\frac{1}{51}$, because one card is now missing from the deck. Strong independence assumptions state that no previous features influence the other features. In the case of pulling cards, these assumptions would lead to the conclusion that the probability of pulling a specific card is $\frac{1}{52}$ every time. Strong independence assumptions introduce simplicity that drastically reduce the time it takes to both create and use NB. However, they do introduce accuracy issues for the model.

NB modifies itself by changing the values of $p(y_i|x_1)$, $p(y_i|x_2)$ etc. each time it observes new training data. Each time it sees a particular feature associated with a classification, it increases the probability of that feature being associated with that class. The inverse happens when a particular feature is not associated with a classification and the probability is decreased for that term. NB then uses the naive Bayes' Theorem to give a probability that something belongs to a particular class.

Logistic regression (LR) is a machine learning model that uses statistics to predict the probability that something will belong to a particular class. LR is based on the logit function, also known as the log-odds function. The logit function is $log(\frac{p}{1-p})$where $p$represents probability. Using the same notation as NB, this can be written as

$$log(\frac{p(X|Y)}{1-p(X|Y)}) \tag{3}$$

where $X$is the matrix of features from the phenomenon being observed, and $Y$is the matrix of discrete classifications that the phenomenon could belong to. $p(X|Y)$is defined as

$$p(X|Y) = \frac{1}{1-e^{-f(x)}} \tag{4}$$

where $e$is Euler's number and $f(x)$is the function of the features and their weights (Subasi, 2019). The value of the logit function corresponds to the probability of a given phenomenon belonging to a certain class. When the model analyzes sets of features in the data that correspond to a particular class, the weights for those features are correspondingly increased (Subasi, 2019; Mangasarian et al, 1995). Conversely, the weights are decreased for a set of features that does not correspond to a particular classification. Applying the logarithmic function to the probabilities reduces the impact of outliers on the data compared to other more linear models (Subasi, 2019). LR easily incorporates new training data because the logarithm normalizes new probabilities (Mangasarian et al., 1995; Chou et al. 2001). LR is therefore easily trained even after the model is complete, unlike many other types of models.

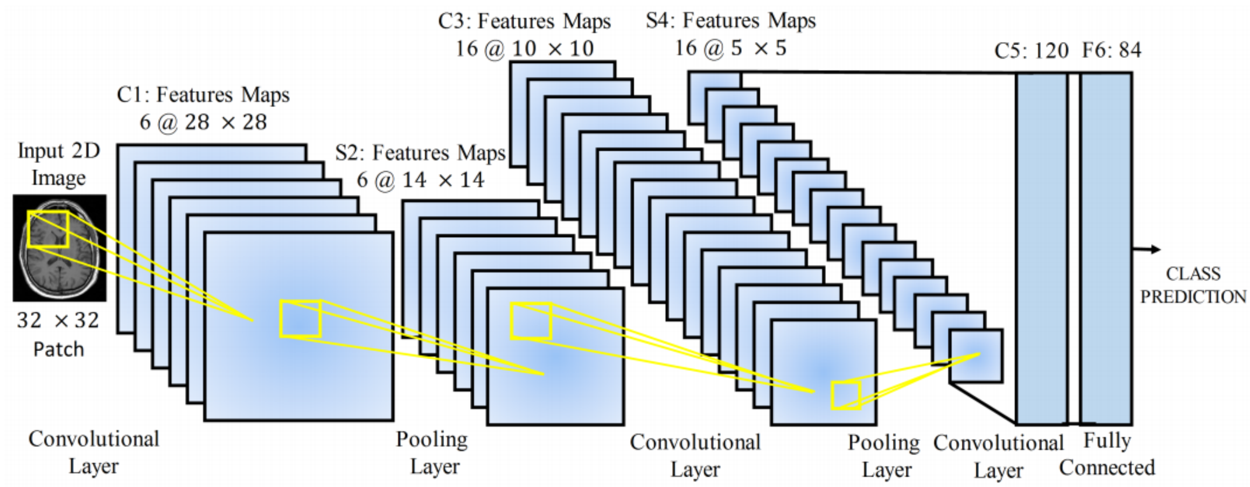


Figure 9: A visual representation of a Convolutional Neural Network for classifying images (Anwar et al., 2018, 11).

Convolutional neural networks are a type of ANN specialized for pattern recognition in images. A typical CNN consists of an input layer, output layer, and many hidden layers. Figure 9 shows the architecture of a typical CNN. The input layer contains the image to be analyzed. The output of a CNN is typically a list of probabilities that the image contains certain features. A CNN's hidden layer is complex; it contains a number of "convolution" and "pooling" layers. In the convolutional layers, matrices of weights, known as filters, are applied to regions of the image. Each filter corresponds to a particular pattern that may appear in an image, such as a straight vertical line or a circle. The application of a filter across an image gathers information about where that feature appears in the image. Each filter then creates a feature map of where the corresponding feature appears in the image. Once the convolutional layer has applied each filter to the image, the feature maps generated are reduced in size. This process of reducing feature map size is known as pooling and involves keeping only the most prominent features. The process of convolving and pooling is repeated multiple times, with increasingly sophisticated filters to extract more complex features. After multiple iterations, the final feature maps are passed to a fully-connected layer. The fully-connected layer processes the extracted features together to classify the image. For example, if the features are a car tire, car headlights, and a car door, the fully-connected layer would classify the image as a car. A CNN can classify images by extracting features with convolution then analyzing them with a fully-connected layer.

2.2.2 Artificial Intelligence in Otology

While the use of AI in the otology is uncommon, some ML models have already been implemented. One example includes the use of an AI to diagnose patients with vertigo. Diagnosing vertigo is a difficult task because symptoms of vertigo are shared among a variety of diseases (Joutsijoki, 2013). Despite this, Ver, a machine learning system, used medical data to diagnose patients thought to be suffering from vertigo. This machine learning system achieved a 79% success rate. Humans, meanwhile, reached a success rate of 68% (Joutsijoki, 2013). Diagnosing vertigo is not as straightforward as other diseases as it shares symptoms with ear infections, damaged inner ear, and even tinnitus (Joutsijoki, 2013). The high accuracy of the AI compared to humans supports the use of AI in otology.

Additionally, in another study, researcher Michelle Viscaino investigated how accurately AI selected the proper ear wax plug for patients. After training the AI model with 720 images of data from 180 patients, AI correctly selected the proper earwax plug 93.9% of the time (Viscaino, 2020). When doctors were asked to complete the same task, they correctly identified the proper earwax plug 95.3% of the time. AI performed similarly to the doctors while being significantly more efficient (Viscaino, 2020). Although AI's use in otology is nascent, these studies showcase the potential of AI.

2.3 Research and Goals of the UHZ Otology Department

To better understand the research goals and objectives of the UHZ Otology department, our team worked closely with our sponsor, Dr. Ivo Dobrev, and his team. The UHZ Otology department is involved in a variety of project areas, spanning from clinical to research applications. Dr. Dobrev's research primarily focuses on improving measurement systems for acoustic response in the middle ear. The results of this research will be applied in clinical diagnostic procedures of hearing loss, as well as the design of hearing aids. In addition to this work, there are members of Dr. Dobrev's team that focus on the clinical aspects of otology, such as surgery. Dr. Christof Röösli, another individual with whom we worked closely, is involved in surgeries of the inner ear to remove VS tumors. The goal of this project was to survey the areas of the otologists' work that would benefit from the use of AI solutions. This report focuses on ways AI can assist the diagnosis and classification of VS tumors in the inner ear.

Chapter 3: Methodology

The goal of this project was to discover the types of artificial intelligence (AI) tools that can assist the University Hospital Zürich (UHZ) otology department's research. We gathered quantitative and qualitative data from interviews, a survey, and the websites of companies that specialized in AI systems. Our team used these 3 data collection mediums to gather data to answer the following two research questions:

1. In what ways can artificial intelligence tools improve surgical operations and medical diagnosis in the field of otology?

2. How have artificial intelligence tools been implemented in the Swiss medical field?

The data collected for these objectives was compiled and used in the formation of our guidebook for the UHZ Otologists. This guidebook informs the otologists on the fundamentals of different AI systems, along with some of the AI companies that could potentially assist them in their current and future research.

3.1 How can artificial intelligence tools improve Otology?

Our first objective was to find ways in which AI can improve the medical diagnoses and surgical operations within otology. We compiled information through nine interviews and one survey with members of Dr. Dobrev's team. We wanted to understand the projects the researchers were involved in. The data collected helped us identify the potential areas for AI applications. We looked to gain insight on the challenges faced by the otologists so we could better recommend AI solutions.

To accomplish this objective, we conducted a series of interviews and a survey with Dr. Dobrev's team. We asked the otologists to participate in an initial survey before their first interview. This survey asked about their previous exposure to AI and where they believed AI could improve their work. We then completed the first round of interviews with seven different otologists who completed our survey. For the interviews, two members of the team led the interview while the other two members of the team took notes. The notes were then used as references for the rest of the project. Throughout these interviews we gathered more information on the researchers' current projects. We then conducted independent research into each project to decide which ones would be the best fit for AI solutions. The projects with potential to be assisted by AI tools prompted our team to conduct a secondary interview with the otologist. In this second interview, we sought to gain a deep understanding of the project, and information about potential uses of AI. We remained in contact with the otologists over email to ask any additional questions that emerged over the course of the project.

We chose to use interviews because they allowed us to find out what was important to industry professionals in AI and in otology. Once we scheduled these interviews, we only needed about an hour to meet with each member of the team. With their consent, the interviews were recorded and referenced during our data synthesis. Another reason we chose to use interviews is to get a more in depth look within each individual project. With interviews we were able to go in depth about the specific problems faced by each otologist. This gave a deeper understanding that allowed us to continually narrow the scope of the project down to just a few potential applications.

Our team also interviewed a local AI expert. In this interview we sought to gain understanding of how AI could best be implemented into the UHZ otologists' work. This interview was allotted a two-hour time slot and was conducted over Zoom. The first part of the interview consisted of our team using the synthesized data from our interviews with the otologists to inform the expert on the problems the otologists are facing. In the second part of the interview we asked questions regarding the feasibility of AI solutions for various projects, as well as the most effective methods of implementation for these projects. This local interview allowed us to better recommend the types of AI that would work effectively within otology and how they could be created and used.

3.2 How have artificial intelligence tools been implemented in the Swiss medical field?

Our second objective was to discover the ways AI has been implemented in the Swiss medical field. The data collected for this objective was used to help inform the UHZ researchers of medical products and technologies that required AI. This information allows the researchers to understand AI's capabilities and the ways in which it could be adapted to suit their research. Additionally, this information was used to recommend AI companies in our deliverable guidebook. The AI companies included in the guidebook are those that we found the otologists may be interested in contacting for future research projects.

Accomplishing this objective required our team to conduct internet research and interviews of AI companies. Using the internet, our team researched AI companies in the Zürich area. When we began our search, we primarily looked at any company that incorporated any AI into their product or technological solution. As we compiled the results from the interviews with the otologists, it became easier to narrow our search to companies that specialized in medical research or life sciences. When our team found a company that could produce helpful tools for the otologists at UHZ, we explored their company website for more information. Often, companies will showcase important information about their products and services on their websites. Using this, we gathered consumer level information about the company's services and how they might best suit Dr. Dobrev's team. To determine which companies to interview, we constructed a set of questions that allowed us to make a comparable data set. These questions are:

1. Is this company based in Zürich?
2. Has AI been incorporated into their products?
3. Does this company have experience in the medical domain?
4. Are there any current uses of their products that have been used in the medical domain?
5. What type of AI solution does this company provide?
6. Is there a way in which "X" company's solution could be adapted to suit the otologists?

Based on the answers to these questions, we considered if each company's work could be applied to the UHZ otology department's research. If so, we requested an interview to gain more information about their products and solutions that could not be found on their company website. If we found that a company answered four or more of these questions properly and completely, we reached out for an interview. The email script used for requesting an interview can be found in Appendix [X]. Each email included a section that was tailored to the specific company which indicated we understood the general idea of their product and were looking to acquire more information.

Our team sent out 16 emails to companies that we believed had products that could be adapted for the otologists at UHZ. Unfortunately, contacting these companies for interviews proved to be difficult, as only two of the 16 responded to our emails. Nevertheless, we conducted semi-structured interviews with Caresomma, a medical imaging company, and KNIME, a data analytics software company. We saw semi-structured interviews as the most effective data collection method as it allowed our team to receive answers to the structured questions, while also giving us the flexibility to inquire information on other related topics (Gill, 2008). Both interviews were allotted a one-hour time slot and were conducted via Zoom. For each interview, at least two members were present. We also recorded the meeting using Zoom so we could reference it while compiling information. The goal of the interviews with the AI developers was to gain a better understanding of their company's involvement with AI tools and uncover the ways in which their work could assist the otologists in their research. While our team was limited to the number of interviews we conducted, many of the companies we attempted to contact were included in our guidebook deliverable.

Despite setbacks, we were able to answer our research questions: “in what ways can artificial intelligence tools improve surgical operations and medical diagnosis in the field of otology?” and “how have artificial intelligence tools been implemented in the Swiss medical field?” The information from these research questions was synthesized and then compiled into a guidebook for the otologists at UHZ. The guidebook included information about where AI could be implemented into the otologists’ work, how such AI could be implemented, and what AI companies would be best to contact for that work. The delivery of this guidebook accomplished the primary project goal: to identify the most promising uses of AI systems in otological research, and to inform the members of Dr. Dobrev's team on such applications.

Chapter 4: Findings

This section discusses vestibular schwannoma (VS) and how artificial intelligence (AI) can be used to classify these tumors as likely to grow or not. We discuss the factors affecting the viability of a tumor classification AI such as the amount of training data and the correlation between image features and tumor growth. We then detail possible AI models that have strong classification abilities. Finally, we evaluate each model's strengths and weaknesses.

4.1 Tumor Diagnosis

VS tumors (discussed previously in section 2.1) are often diagnosed using magnetic resonance imaging (MRI) (C. Röösli, personal communication, September 28, 2020). Diagnosing the tumor once the MRI has been taken is straightforward, as the tumor appears as an unusual bright spot in the scan. While identifying a tumor's presence is simple, predicting if the mass is likely to grow or not is more difficult. Currently, to identify tumor growth, doctors take multiple scans of a patient over a period of several months. By successively measuring the diameter of the tumor, doctors can visualize the tumor's size in relation to the first scan. If the scans indicate that the diameter of the tumor is increasing, doctors will take MRI scans more frequently. If the tumor , doctors perform an invasive tumor removal surgery. Since surgeons are operating on nerves responsible for interacting with the brain, there is a high risk of complications (C. Röösli, personal communication, September 14, 2020). Unfortunately, once a tumor is identified, there is no reliable way to predict if the tumor will grow. Until multiple scans are taken, doctors and the patient have very limited knowledge on the likelihood of the tumor needing removal.

The creation of an accurate machine learning model for diagnosing VS would bring multiple benefits to the otologists at UHZ. A model that accurately predicts whether a tumor will grow could assist a doctor in recommending patient treatment options. New patients whose tumors are predicted to grow could undergo surgery earlier than with the typical diagnostic process. At the same time, patients whose tumors are predicted not to grow can avoid undergoing the highly invasive tumor removal procedure. The use of an AI to assist in diagnosis would not replace the standard method of taking multiple images months or years apart. Rather, the AI would assist in predicting tumor growth. The otologists at UHZ could still measure tumor growth that was not predicted by the AI.

The MRI scans gathered by UHZ are suitable for pattern recognition because of the large maintained data set and consistent properties of each image. One of the main concerns found when creating a medical-based AI revolves around the amount of training data available. As mentioned previously, the researchers at UHZ and other local hospitals have maintained an extensive dataset of MRI scans of VS tumors (C. Röösli, personal communication, October 7, 2020). The file of each patient contains information on when each scan was taken, along with the progression of the tumor growth. Images of VS tumors between patients contain many similar traits. The two most important traits include the location and the shape of the tumor. Since VS only grows on the auditory nerve, this decreases the amount of data the AI needs to internally process, allowing it to focus more specifically on the tumor. Also, according to Dr. Röösli, the structure of VS tumors are found to be relatively similar across many samples. Once again, this is beneficial to the AI model because it makes the pattern recognition process much simpler. The more commonalities found between series of scans, the easier AI can identify patterns to determine the chance of tumor growth.

There are many possible features that could contribute to VS tumor growth. However, the exact features that cause growth are still unknown (C. Röösli, personal communication, September 28, 2020). Using AI, it may be possible to identify factors that lead to an increased chance of tumor growth. It could also be the case that MRI images, while useful for diagnosis, do not contain any features that indicate growth. From prior studies of breast cancer, there are some factors that seem to impact the chance of a tumor being malignant and growing. The aforementioned studies gathered information on the following aspects of tumors found in the breast: density, shape, texture, symmetry, and cystic properties (Chou et al., 2001, Mansagarian et. al 1995). According to the results of their studies, the researchers believed that these were five of the primary factors that affected the tumor's likelihood of being malignant. These breast cancer studies used ultrasound images, but many of the characteristics that affect tumor malignancy are also features found in MRI scans. It could be possible for images of VS tumor images to correlate with tumor growth in the same way as breast cancer images.

Pattern recognition is not possible without proper amounts of training data. According to Dr. Christof Röösli (personal communication, October 7, 2020), UHZ has received approximately fifty new patients each year diagnosed with VS since 2017. For each of these new patients, a scan is taken at the initial diagnosis, and then again six months afterwards. The time between scans is then increased to a year if the tumor is not growing. The database at UHZ currently has images from 255 patients diagnosed with VS; each of whom had multiple MRI scans taken to monitor the growth of the tumor (C. Röösli, personal communication, October 7, 2020). Before the UHZ otologists train an AI with their image data, the data must be anonymized. Pending ethical approval from UHZ, it would be approximately a four-month process to anonymize the images in these databases (C. Röösli, personal communication, October 7, 2020). Once completed, the AI can be trained.

There are multiple methodologies for developing an AI system that may be effective in diagnosing VS tumors. The otologists at UHZ could train an AI to classify tumors based on single images only, or multiple images with timestamps associated.

- **Single Images.** Training an AI on single MRI images would allow it to classify a tumor during a new patient's first scan. However, it is not known if it is possible to predict whether a vestibular schwannoma tumor will grow based on a single scan. Therefore, this method may not be accurate. The training data for this method would need to remove duplicate scans from the same patient to avoid overfitting.

- **Multiple Images.** Training an AI on sets of multiple MRI images would prevent it from classifying a tumor from a new patient's first scan. By giving the AI information about a tumor at two different points in time, the AI may be able to detect patterns of change that commonly lead to growth. For example, it may be that tumors often become more cystic before growing, making cystic properties an important feature for classifying growing tumors. This method may result in higher accuracy but would be more complex to implement due to the increased inputs. The training data would not face as many issues of overfitting as single image training.

In addition to choosing a single or multiple image approach, otologists at UHZ would also need to consider including medical history in a tumor classification AI.

- **Analyzing Images and Medical History.** A patient's medical history and information, such as weight, age, and past conditions, may be important factors in determining tumor growth. By including these factors in the AI system, it may be possible to detect patterns of medical information that correspond with tumor growth. For example, it may be that age correlates strongly with tumor growth, even when image patterns do not. This method would significantly increase the complexity of training the AI systems, as images would have to be passed as input with their associated medical history. The training data would need to use images from patients with a variety of medical histories and backgrounds to avoid overfitting. It would also require associating any new patients' scans with their medical history.

- **Analyzing Images Only.** While medical history may be an important factor in predicting tumor growth, it is also possible that the MRI images themselves could be used to predict tumor growth. By only using images as input to the AI system, much of the time needed to construct a training data set would be reduced. Furthermore, it would reduce the complexity of the AI system. However, it is possible that images do not predict tumor growth as well as medical history, and in that case this method would lose accuracy.

4.2 Implementation

According to our research and the AI experts we spoke to, ensembling multiple AI models together will give the best results in implementation. Ensembling is the process of using multiple AI models in conjunction with each other to create an output. Once the models are created, the otologists can either use the model with the highest accuracy or merge two similar models to create an even more accurate model. To this end, we identified the advantages and disadvantages of the most promising classifying models. These models are the naive Bayesian (NB) classifier, the logistic regression (LR) classifier, and the convolutional neural network (CNN). We also found that using an autoencoder in conjunction with another AI model would reduce computational time and increase accuracy.

According to one of the AI experts we spoke to, an autoencoder for the MRI images would break down the image data into the most important features. An autoencoder is a type of AI that breaks down an image into its most defining features, such that the original image can be reconstructed using only the significant features. Autoencoders reduce the total number of features down to simply the important ones. This would reduce computational time for the AI models used after the autoencoder and reduce the impact of unimportant image features on the AI's result. More relevant data increases the accuracy of AI models that can then be used to predict the likelihood of a VS tumor to grow.

The naive Bayes classifier is a type of AI model that uses the Bayes Algorithm to predict the likelihood of an event based on prior data. However, NB assumes that all features of the data are completely independent of each other (Frank et al, 2002; Xhemali et al, 2009). NB is very quick and very easy to set up and implement for use because of the simplicity of the model and the algorithm it runs on (Frank et al, 2002). The model does not need as much training data as other models and is not heavily affected by irrelevant features in the training data (Xhemali et al, 2009). Using NB, the otologists can use all the features from patient medical data to train the AI. NB will largely ignore any of the inputs that do not have an effect on the outcome because of its assumption that all data features are entirely independent which is almost never true however, leading to low accuracy for the model. This speed comes at the cost of overall accuracy. NB would be able to take the data from the image and the patient medical history and use it to quickly predict the likelihood of tumor growth.

Logistic regression models use a sigmoid function to predict outcomes based on training data. LR excels at predicting target variables that are easily classified into specific categories with a given confidence. LR has already been used to successfully predict whether breast cancer tumors are likely to grow or not (Mangasarian et al., 1995). Like the NB model, LR is relatively easy to set up and implement (Chou et al., 2001). LR excels at classifying objects into specific categories. Since the otologists at UHZ need to classify the tumor as either likely to grow or not, LR should not have an issue with the target variable. The model is often written to give a confidence rating when placing an input into a category. This rating could be converted into a percentage and used to interpret the likelihood that a tumor will grow. However, LR is slower to set up than NB, but it does produce more accurate results (Frank et al., 2009). It also requires more data cleaning and pruning compared to similar models. Without this, LR is prone to overfitting.

Convolutional neural networks could be used to achieve high accuracy results in image analysis. CNNs contain a number of convolutional layers which can identify subtle patterns in MRI images (Sarraf et al., 2016). By analyzing these patterns or features with a fully-connected layer, the CNN may be able to classify a tumor as likely to grow or not. CNNs have already been implemented in a number of medical applications, such as breast cancer diagnosis (Charan et al, 2018) and diagnosis of Alzheimer's disease (Anwar et al., 2018, 11). CNNs typically have more accuracy than LR and NB models (Sunarya et al., 2019; Ayer et al., 2010). However, CNNs also require a larger set of training data than LR or NB models to achieve high accuracy. In a study by the Massachusetts General Hospital, CNNs were found to achieve 95% accuracy when trained on a set of 200 CT scan images (Cho et al., 2015). Since UHZ has at least 255 unique patients each year, there is a reasonable amount of training data for VS. Therefore, it is possible for the otologists at UHZ to achieve high accuracy with a CNN that analyzes MRI images of VS. The increased complexity of CNNs may necessitate more time to implement and train.

There are several AI models that could be used to classify VS tumors. These models could be implemented into the UHZ otology department either by the researchers themselves or by AI developers the otologists choose to work with. Our deliverable guidebook contains more information about each type of AI model and includes tutorials for how to create some of the simpler models, such as the NB model. If the otologists wish to work with an outside developer to implement AI into their work, our guidebook also contains several AI companies that work in similar fields. A general overview and the contact information of each company is provided within the guidebook. This gives the otologists the choice to either try to implement these models themselves, or work with other AI developers to use AI for VS tumor classification.

Chapter 5: Conclusions

This section discusses our final conclusions and recommendations for Dr. Dobrev's team regarding which AI models to implement, further avenues of research, and which AI companies in Switzerland may be able to assist in this research. This includes the most promising AI companies that we recommend the University Hospital Zürich (UHZ) contact and the ways these companies would benefit their research. We also discuss many types of AI, including ensembling, autoencoding, and logistical regression, and the ways in which each of these are used. We discuss any limitations that will require further research before these solutions are implemented.

5.1 Recommendations

We recommend using patient medical history and MRI image data to train an AI to classify vestibular schwannoma tumor growth. The factors contributing to VS tumor growth are not yet fully understood. Aspects of the patient's medical history could be contributing factors to the growth of the tumor. Care must be taken to ensure that this data remains anonymous, which can add additional time to the data cleaning stage of development. We think the benefits of including this data outweighs the costs. The data contained in patients' medical history could allow otologists to identify additional risk factors that contribute to tumor growth. Using this data to train the AI could allow for a more accurate model.

We recommend ensembling multiple different AI models with an autoencoder to identify tumor growth. The autoencoder reduces the number of features in the MRI image to only the most important ones. Three-dimensional MRI images may be too complex to process with multiple AI models without some form of compression. An autoencoder improves the performance of successive models, especially logistic regression (LR). Ensembling is a process of using multiple AI models in parallel and combining their results. By combining the results of the models - or selecting the most accurate model - otologists can achieve more accurate results.

We recommend the otologists at UHZ use a naive Bayes (NB) Classifier as the first model created to identify tumor growth. Choosing this model first would allow the otologists to check the feasibility of the project without investing as much time. The advantages of NB are that it is fast, easy to implement, ignores irrelevant features, and does not require as much training data as other models. However, NB makes significant assumptions about the data, which can lead to accuracy issues. If NB can predict tumor growth, then the otologists will know that the prediction is possible. Then, more accurate models can follow NB to improve upon the predictions. NB is best used as a test model to ensure that it is possible for VS tumor growth to be predicted using AI, rather than in clinical implementation.

We recommend the otologists at UHZ use a logistic regression (LR) model if the naive Bayes model is able to classify tumor malignancy but is not accurate enough to be used clinically. LR is much faster to create and implement compared to convolutional neural networks (CNN), but is slower than NB. LR has been used to predict tumor growth in the past (Chou et al., 2001, Mangasarian et al., 1995). The advantages of LR include quick setup, effective results, and the ability to train the model as it is used. The disadvantages of LR include its low accuracy compared to CNN, and that it is prone to overfitting if the data is not varied. LR serves as an intermediate model between CNN and NB. If LR is accurate enough to be used clinically, there is little reason to advance further to CNN.

We recommend the otologists at UHZ use a convolutional neural network (CNN) if the logistic regression model is not accurate enough to be used clinically. A CNN is an AI model that takes longer than others to implement but is very powerful. CNNs have especially strong pattern recognition, which can make them accurate enough for clinical use. CNNs have been used in medical research before, with successful results. The disadvantages of CNNs include that they are time-intensive and costly to implement, that they require more training data than other AI's, and that they are prone to overfitting if the data is not varied enough. While CNNs are very powerful, they may be more than the otologists need. This model should be used if LR fails to predict tumor growth or is not accurate enough to be used clinically.

We recommend the otologists at UHZ investigate the monetary and time cost of implementing an AI before pursuing machine learning solutions. One area that was not investigated is the monetary cost associated with both creating and training an AI system. Preliminary research has indicated that AI systems can vary in price and the true cost can only be determined once an estimate is conducted. With regards to UHZ otological research, there are two types of AI software that can be created. If a custom AI solution is found necessary, expenses can range anywhere between $6,000 - $300,000 (Elmousalami, 2019). Alternatively, if the otology department is interested in modifying a pre-existing AI, costs can be as high as $40,000. Prices of AI solutions can vary based on their complexity, feasibility and trainability. We recommend that UHZ pursue an AI software that is custom built for their needs since the research being conducted is novel, and there are currently very few AI tools that focus in otology.

Constructing AI models that are accurate and effective requires many months and sometimes even years. The more specific and complex the software, the more time necessary to ensure that it can be adequately trained and tested. It can take a significant amount of time to change data sets meant for human consumption into data sets meant for AI and Machine Learning algorithms (Bosch, Crnkovic & Olsson, 2020). While having a large data set is important for training the AI, the data must be preprocessed to ensure that the machine can properly understand the content. Additionally, for tumor classification AI, the data used in training must be anonymized and requires ethical approval from the UHZ Hospital. According to Dr. Röösli, this process alone can take up to four months. Finally, as with any software, it can take time to test and retest the model to ensure that it is working properly. Most AI systems are not fully automated, and they require some degree of human input. According to a study done on medical AIs at Cornell University, human expert input regarding system improvement is supplied iteratively (Holzinger, Biemann & et al., 2017). Finding an expert who understands both the inner workings of AIs hidden layer and tumor classification can take time and result in delays to the model's creation.

We recommend investigating other methods of problem solving in addition to machine learning solutions. While AI is a powerful solution to many problems, oftentimes it is not the most optimal solution as it is a complex process to program and train it. For simple problems, creating an AI system may be unnecessary as traditional programmatic methods can accomplish the same goal in fewer lines of code (O'Mahony et al., 2019). AI models excel in areas where there are significant amounts of viable training data. From our research, specifically in the work being done in electrocochleography surgery sensors, sufficient training data does not currently exist. Here, more simple tools that can monitor the EcochG curve may be more useful for the otologists. For example, EcochG sensors can incorporate simple programs to track the response of the curve generated to ensure that it is operating within an acceptable threshold. AI excels in solving problems that require pattern recognition and identification. Not all the work being completed by UHZ otologists requires pattern recognition.

5.2 Discussions

We were, of course, limited by the COVID-19 pandemic. This circumstance prevented us from developing this project in and around University Hospital Zürich, and as such, necessitated flexible solutions. We did run into trouble when contacting AI companies directly. Most are located close to UHZ and would have been easily accessible to us had we been in Zürich. Due to the nature of this time period, we could only contact them by email. This was not always easy, as many of them have generic email addresses for general queries, formatted as info@companyname.com. Despite this, we did receive a useful number of responses from these companies and were able to schedule interviews with a few of them.

However, there were also benefits of working remotely for this project. For example, we could schedule interviews and meetings more flexibly, and accommodate the scheduling needs of those we interviewed. Also, as we did not have to meet up in person to work on the project, our team could work together without as much limitation. This meant that we could contribute to the project at any time we decided to and helped our work considerably.

We met with several otologists from UHZ and investigated a few projects in addition to vestibular schwannoma tumor classification. Our sponsor, Dr. Ivo Dobrev, is working on bone conduction research with cadaver skulls. This research measures the propagation of sound through bone when emitted by various hearing devices, such as hearing aids. Based on the measurements, Dr. Dobrev can evaluate the hearing devices - finding optimal placement locations and problems at specific frequencies. He conducts this research for both independent companies and the Swiss government. The research requires manual identification of patterns in three-dimensional graphs of vibrations in the skull, so it takes a long time to perform. We considered the possibility of using an AI to automatically detect patterns in this data.

Unfortunately, we found that due to a lack of training data AI would not be implementable in Dr. Dobrev's research at this time. As more research is done and more training data is collected, it could potentially be a new avenue of research. This could lead to applications in using AI to identify the relationship between sound frequency and the vibrations in the skull.

Other otologists we met with were working on using electrocochleography (ECochG) to improve cochlear implant surgeries. We discussed the possibility of using AI to monitor electrical signals produced by the cochlea during cochlear implant surgery. When the cochlea is damaged, the signals it produces when stimulated are changed. These changes and signals are dependent on individual ear geometry, hearing loss, and the position of the measuring electrode. An AI could be trained on ECochG data from many patients, then used to examine the signals produced by a cochlea in real time during a surgery. This would give immediate feedback to a surgeon when they have made a mistake. Currently the surgeons must wait at least a month until they receive feedback on how well the surgery was performed. With this, surgeons would be able to feel what caused them to make a mistake and improve their technique to make fewer mistakes in the future. This would also lead to better overall surgery outcomes. If the AI detects cochlear trauma the surgeon can stop, reevaluate, and decide to either try again or stop the surgery.

While we believe this project is well suited to AI, we found a few problems that should be addressed before AI can be applied in this field. First, training data is being gathered slowly, and there is not a large available database at the moment. More data would need to be gathered before AI could be trained to recognize cochlear trauma. Furthermore, the effects of trauma on ECochG signals is not fully understood yet. This must be further investigated before a data set could be properly prepared for an AI. The signal processing would also have to be done in real time. This brings in the issue of computational time and adds restrictions to what models can be applied.

When enough training data has been prepared, the otologists will want to look for a classifier. The AI would want to classify the signals it receives as either "trauma" or "not trauma." This is similar to the tumor diagnostics AI, which classifies a tumor as either likely to grow or not. Most of the same classifiers can therefore be used for classifying ECochG signals. While NB and potentially LR could be fast enough to perform in real time. CNNs are likely too slow to be used for this application. Other classifiers may want to be investigated more thoroughly for this project as well, such as support vector machines or K-nearest neighbors models.

5.3 Conclusions

We have found that AI could be successfully implemented for vestibular schwannoma (VS) tumor diagnoses performed by the UHZ otology department. A successful AI could predict the likelihood of malignancy of VS. The otologists and surgeons can identify growing tumors with fewer scans and remove dangerous masses before they result in a serious problem. VS tumors that are small and unlikely to grow would not need to be monitored as frequently as they are now, where the likelihood of growth is unknown. Small tumors that are unlikely to grow do not need to be monitored as closely. Instead of having yearly scans, these patients could come in every two years or even later. This would decrease the number of unnecessary tumor resection surgeries performed. They could instead be monitored using the AI model in conjunction with the scanning procedure.

The AI technology developed for UHZ hospital can be expanded to the surrounding area's hospitals. Other otology and neurology departments could use the created AI to improve the diagnosis of VS. The autoencoder used in the development of the AI ensemble could be adapted to work with MRI images of other ailments. This could lead to faster implementation of AI within the diagnostic processes. Less time would be spent creating a new autoencoder for the process.

Predicting VS tumor growth could lead to a better understanding of the factors that cause this tumor growth. Coupled with an examination of the factors contained within a patient's medical file, this could lead to finding definitive risk factors for the growth of VS. A stronger understanding of these factors is important in developing more effective treatments and preventative measures.

Our project set out to find where AI could be implemented within the work of the otologists at UHZ. We researched both the fields of otology and machine learning to find a possible intersection between the two. We found that AI can be an effective tool in the classification of VS tumors seen in MRI images. In investigating this aspect of the otologist's work, we found viable AI solutions that could lead to the successful implementation of AI within otology. According to our research, if the otologists ensemble several different machine learning models with an autoencoder, they will likely be able to predict the malignancy of a tumor based on the scans taken at initial diagnosis. This solution could be shared outside of the University Hospital Zürich to improve the diagnosis of VS at several other hospitals.

Bibliography

Asanka, D. (2020, April 14). *Implement Artificial Neural Networks (ANNs) in SQL Server.* SQLShack. Retrieved September 30, 2020, from https://www.sqlshack.com/implement-artificial-neural-networks-anns-in-sql-server/.

Ayer, T., Chhatwal, J., Alagoz, O., Kahn Jr, C. E., Woods, R. W., & Burnside, E. S. (2010). Comparison of logistic regression and artificial neural network models in breast cancer risk estimation. *Radiographics*, *30*(1), 13-22.

Bahrammirzaee, A. (2010, June 20). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, (19), 1165–1195. ResearchGate. 10.1007/s00521-010-0362-z.

Baker, R. R., Lott, S., Cherry, J., & Hoopes, J. (1965). Malignant tumors of the head and neck. Maryland Medical Journal, 14(12), 49-50.

Balasubramanian, G. K., Thirunavukkarasu, R., Kalyanasundaram, R. B., & Narendran, G. (2015). Assessment of Hearing Status by Pure Tone Audiogram—An Institutional Study. *International Journal of Otolaryngology and Head & Neck Surgery*, *4*(05), 375.

.

Bance M. (2007). Hearing and aging. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*, *176*(7), 925–927. https://doi.org/10.1503/cmaj.070007.

Berger A. (2002). Magnetic resonance imaging. BMJ (Clinical research ed.), 324(7328), 35. https://doi.org/10.1136/bmj.324.7328.35.

Bosch, J., Crnkovic, I., & Olsson, H. H. (2020). Engineering AI Systems: A Research Agenda. *arXiv preprint arXiv:2001.07522*.

Brinker, T. J., Hekler, A., Utikal, J. S., Grabe, N., Schadendorf, D., Klode, J., Berking, C., Steeb, T., Enk, A. H., & von Kalle, C. (2018). Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review. *Journal of Medical Internet research*, *20*(10), e11936. https://doi.org/10.2196/11936

Charan, S., Khan, M. J., & Khurshid, K. (2018, March). Breast cancer detection in mammograms using convolutional neural network. In *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (pp. 1-5). IEEE.

Chou, Y. H., Tiu, C. M., Hung, G. S., Wu, S. C., Chang, T. Y., & Chiang, H. K. (2001). Stepwise logistic regression analysis of tumor contour features for breast ultrasound diagnosis. Ultrasound in medicine & biology, 27(11), 1493-1498.

Chung, J. H., & Lee, S. H. (2017). Cutting Edge Technologies in Otology Field. Hanyang Medical Reviews, 37(1), 25-29.

Cunningham, L. L., & Tucci, D. L. (2017). Hearing Loss in Adults. The New England journal of medicine, 377(25), 2465–2473. https://doi.org/10.1056/NEJMra1616601.

McCarthy, G. (n.d.). Diagram of the Ear - Online Biology Dictionary. Retrieved October 10, 2020, from https://www.macroevolution.net/diagram-of-the-ear.html

Elmousalami, H. H. (2019). Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction. *arXiv preprint arXiv:1909.11637*.

Frank, E., Hall, M., & Pfahringer, B. (2002, August). Locally weighted naive bayes. In *Proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence* (pp. 249-256). Morgan Kaufmann Publishers Inc.

Greene J, Al-Dhahir MA. Acoustic Neuroma (Vestibular Schwannoma) [Updated 2020 Aug 10]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470177/.

Hamilton W. (2010). Cancer diagnosis in primary care. *The British journal of general practice : the journal of the Royal College of General Practitioners*, *60*(571), 121–128. https://doi.org/10.3399/bjgp10X483175.

Harrel, M., & Bradley, M. (2009). *Data collection Methods: Semi-structured interviews and focus groups.* (). Santa Monica, California: RAND Corporation. Retrieved from https://www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR718.pdf.

Holman, J. A., Drummond, A., Hughes, S. E., & Naylor, G. (2019). Hearing impairment and daily-life fatigue: A qualitative study. *Null, 58*(7), 408-416. doi:10.1080/14992027.2019.1597284.

Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2017). What do we need to build explainable AI systems for the medical domain?. *arXiv preprint arXiv:1712.09923*.

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. (2018). Artificial intelligence in radiology. *Nature reviews. Cancer, 18*(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5.

Institute for Quality and Efficiency in Health Care. (2017, November 30). *Hearing loss and deafness: Normal hearing and impaired hearing*. InformedHealth. Retrieved October 10, 2020, from https://www.ncbi.nlm.nih.gov/books/NBK390300/.

Ismail, K. (2018, October 26). *AI vs. Algorithms: What's the Difference?* CMSWire. Retrieved September 28, 2020, from https://www.cmswire.com/information-management/ai-vs-algorithms-whats-the-difference/.

Kuchling, A. M. (2012, August 16). *Background: Algorithms*. 50 Examples for Teaching Python. Retrieved 10 3, October, from https://fiftyexamples.readthedocs.io/en/latest/algorithms.html.

Lawrence, M. (1966). How we hear doi:10.1001/jama.196.10.931.

Liberman, M. C., Liberman, L. D., & Maison, S. F. (2015). Chronic Conductive Hearing Loss Leads to Cochlear Degeneration. PloS one, 10(11), e0142341. https://doi.org/10.1371/journal.pone.0142341.

Martinez-Devesa, P., Barnes, M. L., & Milford, C. A. (2008). Malignant tumors of the ear and temporal bone: a study of 27 patients and review of their management. Skull base : official journal of North American Skull Base Society ... [et al.], 18(1), 1–8. https://doi.org/10.1055/s-2007-992766.

Mehta, R., Zhu, R., & Cheema, A. (2012). Is noise always bad? Exploring the effects of ambient noise on creative cognition. *Journal of Consumer Research*, *39*(4), 784-799.

Mozaffari, M., Jiang, D., & Tucker, A. S. (2020). Developmental aspects of the tympanic membrane: Shedding light on function and disease. Genesis, 58(3-4), e23348. doi:10.1002/dvg.23348.

National Institute of Deafness and Other Communication Disorders. Noise induced hearing loss. Hearing and Balance. 2014:2.

https://www.nidcd.nih.gov/sites/default/files/Documents/health/hearing/NoiseInducedHearingLoss.pdf.

Nilsson, N. J. (1980). *Principles of Artificial Intelligence* (1st ed.). Elsevier Science. https://www.google.com/books/edition/Principles_of_Artificial_Intelligence/mT-jBQAAQBAJ?hl=en&gbpv=0&kptab=overview.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., ... & Walsh, J. (2019, April). Deep learning vs. traditional computer vision. In *Science and Information Conference* (pp. 128-144). Springer, Cham.

Rajkomar, M.D., A., Dean, Ph.D., J., & Kohane, M.D., Ph.D., I. (2019, April 4). Machine Learning in Medicine. *The New England Journal of Medicine*, *2019*(380), 1347-1358. NEJM. 10.1056/NEJMra1814259.

Ralli, M., Rolesi, R., Anzivino, R., Turchetta, R., & Fetoni, A. R. (2017). Acquired sensorineural hearing loss in children: current research and therapeutic perspectives. Sordità infantile acquisita: stato dell'arte della ricerca e prospettive terapeutiche. Acta otorhinolaryngologica Italica : organo ufficiale della Società italiana di otorinolaringologia e chirurgia cervico-facciale, 37(6), 500–508. https://doi.org/10.14639/0392-100X-1574.

Sarraf, S., Tofighi, G., & Alzheimer's Disease Neuroimaging Initiative. (2016). DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI. *BioRxiv*, 070441.

Stangerup, S. E., & Caye-Thomasen, P. (2012). Epidemiology and natural history of vestibular schwannomas. Journal of Neurological Surgery Part B: Skull Base, 73(S 02), A227.

Steeg, P. Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 12, 895–904 (2006). https://doi.org/10.1038/nm1469.

Stomackin, G., Kidd, S., Jung, T. T., Martin, G. K., & Dong, W. (2019). Effects of tympanic membrane perforation on middle ear transmission in gerbil. Hearing Research, 373, 48-58. doi:10.1016/j.heares.2018.12.005.

Subasi, C. (2019). LOGISTIC REGRESSION CLASSIFIER. Retrieved October 12, 2020, from https://towardsdatascience.com/logistic-regression-classifier-8583e0c3cf9.

Sunarya, P. A., Refianti, R., Mutiara, A. B., & Octaviani, W. (2019). Comparison of Accuracy between Convolutional Neural Networks and Naïve Bayes Classifiers in Sentiment

Analysis on Twitter. *International Journal of Advanced Computer Science and Applications*, *10*(5), 77-86. SemanticScholar.

The Definitive Glossary of Higher Mathematical Jargon — Algorithm. (2019, August 1). Math Vault. Retrieved September 28, 2020, from https://mathvault.ca/math-glossary/#algo.

Hain, Timothy C., M. (2020). *DIAGNOSIS OF ACOUSTIC NEUROMA USING TESTS.* Retrieved September 29, 2020, from https://www.dizziness-and-balance.com/disorders/tumors/acoustic_neuroma/tests.html.

van Beek, Edwin J R, and Eric A Hoffman. "Functional imaging: CT and MRI." Clinics in chest medicine vol. 29,1 (2008): 195-216, vii. doi:10.1016/j.ccm.2007.12.003.

Westman, J. C., & Walters, J. R. (1981). Noise and stress: a comprehensive approach. *Environmental health perspectives*, *41*, 291–309. https://doi.org/10.1289/ehp.8141291.

White HJ, Helwany M, Peterson DC. Anatomy, Head and Neck, Ear Organ of Corti. [Updated 2020 May 22]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK538335/?report=classic.

Xhemali, D., J Hinde, C., & G Stone, R. (2009). Naïve bayes vs. decision trees vs. neural networks in the classification of training web pages. D. XHEMALI, CJ HINDE and Roger G. STONE," Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages", International Journal of Computer Science Issues, IJCSI, Volume 4, Issue 1, pp16-23, September 2009, 4(1).

Yildirim, S. (2020, May 12). Naive Bayes Classifier‑Explained. Retrieved October 12, 2020, from https://towardsdatascience.com/naive-bayes-classifier-explained-50f9723571ed.

Zahnert T. (2011). The differential diagnosis of hearing loss. Deutsches Ärzteblatt international, 108(25), 433–444. https://doi.org/10.3238/arztebl.2011.0433.

Appendix A: Guidebook for Otologists

The purpose of this guidebook is to inform otologists at the University Hospital Zürich (UHZ) on the following three topics: the workings of AI, the educational tools and resources available to learn about AI, and the local AI companies in the Swiss area that could assist in future projects. While this guidebook is specifically tailored to the UHZ Otology Department, the contents can be applied to a variety of other medical fields. This guide is referenced when beginning the process of AI implementation into any project. This guidebook will contain some of the resources needed when looking to implement customized AI solutions. The information found in this document was compiled as part of an Interactive Qualifying Project with the UHZ otology department. This booklet was authored by Alan Healy, Emilia Casagrande, Ashwin Pai, and Holden Snyder.

A.1 Basics of Artificial Intelligence and Machine Learning

A.1.1 The Kinds of Problems Machine Learning Can Solve

Machine learning (ML) is often applied in projects that require data categorization or event prediction. Machines that specialize in classification are called classifiers, and those that specialize in prediction are called regressors. Classifiers output discrete, often categorical, variables while regressors output continues, often numerical, variables. The distinction found between classifier and regressor ML models are dependent on the problem. For example, classifiers have been used to translate written words into computerized words, create facial recognition software, and create spam filters for emails. Applications of regressors include solar power forecasting, weather forecasting, and economic event predicting. Classifiers and regressors form the foundation for many other machine learning models.

AI excels at finding patterns when presented with a large set of training data. Through complex data processing, AI can find patterns that are unrecognizable by humans. This is one reason AI has been applied into many image recognition and classification projects. The more training data an AI receives, the better it performs. Since an AI trains continuously, by constantly referencing its training data, it can perform significantly better than humans when presented with large amounts of data. AI has been used to create facial recognition software, create self-driving cars, and diagnose tumor malignancy.

A.1.2 Problems in Otology that could be Approached with Machine Learning

There is potential for vestibular schwannoma (VS) tumors to be classified as either likely to grow or not using AI. UHZ already has a large dataset of VS tumor images that could be used to train an AI. In fact, previous studies have applied AI in the tumor classification process however, VS tumors have not been the subject of any studies in AI yet. A classification AI that could distinguish between tumors that will and will not grow could reduce the number of unnecessary scans for patients and allow doctors to find growing tumors before the second scan. Classifier AIs like naive Bayes classifiers, logistic regression classifiers, and convolutional neural networks would work well for this application.

AI could also assist in the future use of electrocochleography surgery sensors. However, there does not appear to be enough data currently available in this area to train a model. As more data is collected an AI could be trained. An AI can process potential differences generated by the cochlea. This AI could then distinguish between voltage changes caused by movement of the intracochlear electrode and voltage changes caused by cochlear trauma. A device that incorporates this type of AI technology could assist surgeons in cochlear implant surgeries by alerting the surgeon when cochlear damage is caused in real time. The naive Bayes classifier and logistic regression classifier discussed in section A.2.4 are types of AI that would excel in such tasks. A logistic regression would perform better than naive Bayes because the data is numerical and not categorical. Additionally, it may also prove beneficial to investigate other classifiers such as support vector machines and k-nearest neighbors classifiers.

AI applications can also assist the otologists' research in bone conduction. Here we suggest investigating the use of regression models and unsupervised classifiers. Regression models produce outputs that estimate future values or try to predict future events based on training data. AI could prove useful when predicting how vibrations would propagate across a variety of skulls. Unsupervised classifiers would pool the data together then classify it based on its similarities or differences found by the AI. This would identify patterns within the data and organize it for easier analysis. To this end, it may be helpful to investigate linear regressors, k-nearest neighbors regressors, and cluster analysis classifiers.

A.1.3 Basic Mechanics of Artificial Intelligence

AI models are designed to modify their own algorithms based on their training. As developers pass sample inputs to an AI system, they can evaluate the system's performance. If unsuccessful, the AI system will modify its algorithms to achieve a better result. This process is known as training. With sufficient training, AI systems can become highly accurate in areas such as classification, identification, or prediction (Rajkomar, M.D. et al., 2019, 1348). The general process of creating and using an AI follows this pattern:

- **Understanding the domain.** In order to work with an AI system, it is important to thoroughly understand the problem domain. Factors such as data availability, variety, and complexity will inform the choice of which AI models are reasonable to use.
- **Preparing the data set.** For the system to work properly, data must meet certain criteria. All data used to train the machine must be in the same medium and in a consistent format. It may even be necessary to apply some preprocessing to the data, such as removing empty space around images or blurring high resolution images. This process is known as data cleaning.
- **Discovering patterns.** After creating an AI model and training it on sample data, AI can uncover patterns in the data that were not previously visible.
- **Post processing of discovered patterns.** Developers can analyze and interpret the patterns identified by AI systems to come to conclusions.
- **Putting the results into use.** Communicate the results and possible applications of the patterns found through machine learning.

AI relies heavily on the data used to train it. Therefore, creating an accurate AI system requires careful training from varied data. The process of training involves multiple steps. The first step in training is to acquire data which has already been classified. This may require manual classification of a data set, but once completed, the AI can begin training. AI developers can test it on some portion of the training data. After each input is processed, a "loss function" determines to what extent the algorithm's output was accurate. Based on the output of this loss function, a process called backpropagation is used to adjust the AI's weights and minimize error (Bahrammirzaee, 2010, 1167). A weight is the importance the AI has assigned to a specific feature. If the AI finds strong correlation between a specific input and output, the weight for that input will increase for that output. The specific details of the backpropagation algorithm vary for each AI model. As this sequence of processing data and adjusting weights is repeated, the AI becomes more accurate. More training data usually means more accurate outputs from the AI. However, too much similar training data can lead to accuracy issues.

Training an AI from data that is too similar causes a problem known as overfitting. Overfitting is inaccuracy introduced to the model when trying to classify or predict anything not closely related to the training data. For example, if a developer designed an AI to recognize pictures of dogs, then trained it exclusively on dogs with a solid coat, it would fail to recognize dogs with a spotted coat. The importance of a solid coat was emphasized to the AI, since it was a consistent feature of the training data. With more diverse images of dogs, such overfitting would be less likely to occur. Therefore, diverse training data is necessary to implement accurate machine learning models.

Within AI there is also a difference between supervised and unsupervised learning. Supervised learning is when the inputs of the training data are paired with a desired output value. This shows the AI that a particular set of features correspond to this output. In this case, the data used to train the AI must be modified to include the desired output value for each input. This is often done manually if the training data set is not too large. Rather than looking for patterns from human labelled input data, unsupervised learning looks for probabilities in unlabeled data. Generally, unsupervised learning works by trying to cluster or group unlabeled data according to the similarities it finds in the data. This is often used to identify patterns or trends in data that has not been thoroughly analyzed.

A.1.4 Machine Learning Algorithms that could apply to Otology Research

Naive Bayes Classifier

Overview of naive Bayes

- Works off of Bayes' Theorem
- Makes assumption that all features of the data are independent
- Usually used as a classifier
- Likes categorical data more than numerical data
- Easy to implement
- Very fast computation time
- Can require less training data if independence assumption is close to true
- Has accuracy issues if independence assumption is far from true

Naive Bayesian (NB) models are ML models based on using Bayes' theorem to classify data. Bayes' Theorem is $p(y_i|x_1, x_2 ... x_n) = \frac{p(x_1,x_2...x_n|y_i)*p(y_i)}{p(x_1,x_2...x_n)}$where $x_1, x_2 ... x_n$are the features of the data, and $y_i$is a potential classification. $p(y_i|x_1, x_2 ... x_n)$represents the probability that the features of the data represent something from a particular class y_i, $p(x_1, x_2 ... x_n|y_i)$represents the probability of these features occurring in the data if the classification is assumed to be true. $p(y_i)$ is the independent observed probability of class $y_i$occurring, and $p(x_1, x_2 ... x_n)$is the probability of the features of the data occurring together. The naive Bayes model assumes with Bayes' Theorem that all features from the data are completely independent of each other (Yildirim, 2020). This modifies Bayes' Theorem to look like this$p(y_i|x_1, x_2 ... x_n) = p(y_i|x_1) * p(y_i|x_2) * ... p(y_i|x_n)$. In this case, the probability of a given feature is calculated separately from all other features. A real-world example of this difference would be like the probability of pulling a specific card from a deck of cards. The initial probability of pulling a specific card is $\frac{1}{52}$. If the pulled card is not returned to the deck, the probability of pulling a specific card on the next pull becomes $\frac{1}{51}$, because one card is now missing from the deck. Strong independence assumptions hold that no previous features have an effect on the other features. In the case of pulling cards, these assumptions would lead to the conclusion that the probability of pulling a specific card is $\frac{1}{52}$every time. Strong independence assumptions introduce simplicity that drastically reduce the time it takes to both create and use NB. However, they do introduce accuracy issues for the model.

NB modifies itself by changing the values of $p(y_i|x_1)$, $p(y_i|x_2)$, etc. each time it observes new training data. Each time it sees a particular feature associated with a classification, it increases the probability of that feature being associated with that class. The inverse happens when a particular feature is not associated with a classification and the probability is decreased for that term. NB then uses the naive Bayes' theorem to give a probability that something belongs to a particular class.

The advantages of NB are that it can quickly process data, is easy to implement relative to other AI models, ignores irrelevant features, and requires less training data than other models. However, NB makes core assumptions about the data it is fed that leads to accuracy issues for the model. NB is often used when high accuracy is not needed, when strong independence assumptions are close to true, or as an initial test model.

Learning Materials for naive Bayes models:

- A quick explanation on what NB is and how it works.
  - https://towardsdatascience.com/naive-bayes-classifier-explained-50f9723571ed
- YouTube video on how naive Bayes works with examples of predictions
  - https://www.youtube.com/watch?v=XcwH9JGfZOU
- Scientific article on how naive Bayes can be improved using local weights
  - https://arxiv.org/ftp/arxiv/papers/1212/1212.2487.pdf
- This article gives a quick overview of what naive Bayes is and how it works, and includes a tutorial on how to write one in Python
  - https://dzone.com/articles/naive-bayes-tutorial-naive-bayes-classifier-in-pyt

Logistic Regression Classifier

Overview of logistic regression:

- Uses logit (sigmoid) function to try to classify or predict output
- Can be used to classify or predict
- Highly receptive to new training data
- Wide range of applications
- Relatively simple to implement

Logistic regression (LR) is a machine learning model that uses statistics to predict the probability that something will belong to a particular class. LR is based on the logit function, also known as the log-odds function. The logit function is $log(\frac{p}{1-p})$where $p$represents a probability of an event taking place. Using the same notation as NB, this can be written as $log(\frac{p(X|Y)}{1-p(X|Y)})$where $X$is the matrix of features from the phenomenon being observed, and $Y$is the matrix of discrete classifications that the phenomenon could belong to. $p(X|Y)$is defined as $p(X|Y) = \frac{1}{1-e^{-f(x)}}$where $e$is Euler's number and $f(x)$is the function of the features and their weights (Subasi, 2019). The value of the logit function corresponds to the probability of a given phenomenon belonging to a certain class. When the model analyzes sets of features in the data that correspond to a particular class, the weights for those features are correspondingly increased (Subasi, 2019; Mangasarian et al, 1995). Conversely, the weights are decreased for a set of features that does not correspond to a particular classification. Applying the logarithmic function to the probabilities reduces the impact of outliers on the data compared to other more linear models (Subasi, 2019). LR easily incorporates new training data because the logarithm normalizes new probabilities (Mangasarian et al., 1995; Chou et al. 2001). LR is therefore easily trained even after the model is complete, unlike many other types of models.

LR is faster to create and implement compared to CNN or artificial neural network, but requires more computation time than NB. The advantages of LR are quick setup, effective results, and real time model training. The disadvantages of LR are that it is not as accurate as CNN, and it is prone to overfitting if the data is not varied enough. LR is suitable as a middle ground between the two models. If it is accurate enough to be used clinically there is little reason to advance further to CNN.

Learning materials for logistic regression classifiers

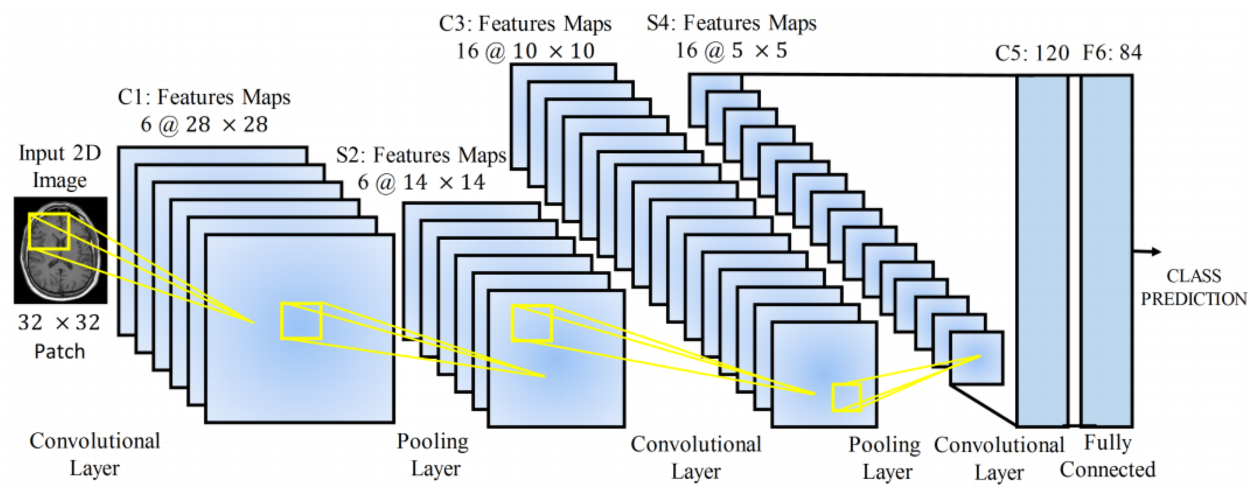
- Article with an explanation of logistic regression. Goes into more detail than we do in this guidebook and covers a good deal of the math.

- https://kambria.io/blog/logistic-regression-for-machine-learning/#:~:text=Logistic%20regression%20is%20a%20classification,either%20a%200%20or%201.
- Simple explanation of logistic regression without the mathematics behind it. This video has follow up videos that go deeper into the inner workings of logistic regression.
  - https://www.youtube.com/watch?v=yIYKR4sgzI8
- This article gives an overview of what logistic regression classification is and how it works. After this overview, it goes through a tutorial of how to write one in Python.
  - https://realpython.com/logistic-regression-python/

Convolutional Neural Networks

Overview of convolutional neural networks

- Type of ANN that specializes in image recognition and classification
- Uses filters to find important features in the data
- Very powerful and accurate
- Requires more training data than other models
- Slow computational time
- Slow, complicated implementation



Convolutional neural networks (CNN) are a type of artificial neural network specialized for pattern recognition in images. A typical CNN consists of an input layer, output layer, and many hidden layers. Figure 9 shows the architecture of a typical CNN. The input layer contains the image to be analyzed. The output of a CNN is typically a list of probabilities that the image contains certain features. A CNN's hidden layer is complex; it contains a number of "convolution" and "pooling" layers. In the convolutional layers, matrices of weights, known as filters, are applied to regions of the image. Each filter corresponds to a particular pattern that may appear in an image, such as a straight vertical line or a circle. The application of a filter across an image gathers information about where that feature appears in the image. Each filter then creates a feature map of where the corresponding feature appears in the image. Once the convolutional layer has applied each filter to the image, the feature maps generated are reduced in size. This process of reducing feature map size is known as pooling and involves keeping only the most prominent features. The process of convolving and pooling is repeated multiple times, with increasingly sophisticated filters to extract more complex features. After multiple iterations, the final feature maps are passed to a fully-connected layer. The fully-connected layer processes the extracted features together to classify the image. For example, if the features are a car tire, car headlights, and a car door, the fully-connected layer would classify the image as a car. A CNN can classify images by extracting features with convolution then analyzing them with a fully-connected layer.

CNNs achieve high accuracy results in image analysis. CNNs contain a number of convolutional layers which can identify subtle patterns in MRI images (Sarraf et al., 2016). By analyzing these patterns or features with a fully-connected layer, the CNN may be able to classify a tumor as likely to grow or not. CNNs have already been implemented in a number of medical applications, such as breast cancer diagnosis (Charan et al, 2018) and diagnosis of Alzheimer's disease (Anwar et al., 2018, 11). CNNs typically have more accuracy than LR and NB models (Sunarya et al., 2019; Ayer et al., 2010). However, CNNs also require a larger set of training data than LR or NB models to achieve high accuracy. In a study by the Massachusetts General Hospital, CNNs were found to achieve 95% accuracy to identify various body parts when trained on a set of 200 CT scan images (Cho et al., 2015). The increased complexity of CNNs may necessitate more time to implement and train.

Learning Resources for convolutional neural networks

- Video Overview of CNNs
  - https://www.youtube.com/watch?v=YRhxdVk_sIs
- Text Guide to CNNs
  - https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
- Journal Article Describing Image Classification with CNNs
  - https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networ
- Playlist of Videos for Machine Learning in Python and Tensorflow
  - https://www.youtube.com/playlist?list=PLQVvvaa0QuDfKTOs3Keq_kaG2P55YRn5v

A.1.5 How to Approach Solving Problems with Machine Learning

There are many ways to implement machine learning into the field of otology. The first is through cloud platforms like the company KNIME (discussed in A.2.1). These cloud computing companies provide a paid subscription service where they take data and run it through their AI services and then return the results.

The second option is for otologists to create the AI themselves. There are many tutorials online, and some are included in this guidebook (an autoencoder for brain MRI images tutorial and a naive Bayes tutorial are included in section A.1.3). There are also many Python libraries available for AI that streamline the process of creating the AI by making the more complicated functions prewritten. One example of this kind of library is Scikit-learn, which is a tool to learn ML in Python. It is an open-source library that includes classification, regression, clusters, and more. It has a tutorial for how to get started with ML included with it. The only required knowledge for this library is a basic understanding of fundamental AI concepts like training data, features, inputs, and outputs.

The third option is to hire a data science or machine learning company to develop the AI with guidance from the otologists. A machine learning company may already have an AI model that they can adapt to the specific needs of otologists. This would reduce the amount of time spent creating an AI and allow the otologists to focus on the research as opposed to the data analysis. This could also be a costly option, and it would be important to anonymize data because of the introduction of a third party.

Learning Resources for Artificial Intelligence:

- Tutorial on how to create a convolutional autoencoder that works on brain MRI images in Python
    - https://www.datacamp.com/community/tutorials/reconstructing-brain-images-deep-learning
- Scikit-learn is a machine learning library in Python with beginner-focused tutorials and many features. Here is their getting started guide.
    - https://scikit-learn.org/stable/getting_started.html

A.2 What companies in Zürich should we contact regarding our ML projects?

A.2.1 KNIME

Summary

KNIME is a data analytics software company that specializes in creating solutions for data analytics problems. KNIME software gives its users the flexibility to create personalized software through its open-source platform. Tumor classification and electrocochleography sensor monitoring are two applications where the KNIME software would excel. It can work in combination with many preexisting databases. This eliminates the need to restructure data to ensure that it is acceptable for the program. Additionally, KNIME software requires minimal programming experience. Many of the data analytics tools are preexisting and are "plug and play". However, for more complex mathematical operations and data processing some programming may be required. But independent code can easily be incorporated into the KNIME workflow.

Notable Publications:

Anomaly detection: KNIME software was used as a predictive analysis tool that monitored the signals generated from machinery to ensure that it was operating at peak performance. A model was trained to detect anomalous signal data, which was used to predict impending breakdowns. This application of KNIME software is notable for otology work in electrocochleographic sensor measurement.

Contact Information:

- Website: https://www.knime.com/
- Email: contact@knime.org

Notable People:

- Dr. Stefan Helfrich is a Project Manager at KNIME with 5+ years of experience with the software. He also specializes in bioinformatics and bioimage analysis which could be beneficial for oncology research and study. We previously conducted an interview with him to learn more about KNIME. He can be contacted at stefan.helfrich@knime.com.

A.2.2 4Quant

Summary:

4Quant is a bioimage analysis company that is looking to bridge the gap in the backbone of medical AI business. They are experienced in a wide range of modalities (PET, CT and MRI) on a variety of conditions from diseases from bone fracture to neuroimaging. BIDAP (Big Image Data Analytics Platform) is 4Quant's core platform incorporates scalable image analysis tools to end-users everywhere. They have worked on analyzing tumors collected from X-rays that are smaller than 5 cm. Additionally, BIDAP program will present the information collected in a statistical report through figures and tables complete with explanations. They have designed the software to ensure that results can be reliably and automatically reproduced. They also offer introductory courses developed at Swiss Federal Institute of Technology on "Quantitive Big Imaging" (http://kmader.github.io/Quantitative-Big-Imaging-2016)

Notable Publications:

4Quant has published many papers on radiomics, a method that extracts a large number of features from medical images using data-characterization algorithms. They have developed a ML solution that can partition a region of interest in images such that each region corresponds to one or more anatomic structures. This has been used in tumor research to understand how the volume of the tumor is changing during treatment. These segmentations can be viewed in real-time and be used to update surgical plans which may assist surgeons who are looking to remove acoustic neuromas from the inner ear.

Contact Information:

- Website: https://4quant.com/
- Address: Heinrichstrasse 249, 8005 Zürich, Switzerland
- Email: info@4Quant.com
- Social Media: @4Quant

A.2.3 ScopeM

Summary:

ScopeM is a research subgroup from the ETH Zürich Hospital. ScopeM is a company that completes data analysis for image based biomedical research. They use ML and Data Analytics models to extract important features from a variety of scans. They are currently invested in developing deep learning techniques for image classification. They are developing tools that automatically provide an assignment of different images to specific classes. For example, from an image the product can determine if the cells present are cancerous or non-cancerous. It is difficult to determine if they are interested in the predictive analysis domain. But they have professionals from biology and bioinformatics that are willing to answer questions through in-person or phone conversations.

Notable Publications:

Currently there are no direct case studies found on ScopeM's website. However, there are a variety of other educational resources accessible through links and research databases.

Contact Information:

- a. Website: http://www.let-your-data-speak.com/
- b. Phone Number: +41 44 633 0759
- c. Email: ida@scopem.ethz.ch
- d. Address: John-von-Neumann-Weg 9, 8093 Zürich

A.2.4 IBM

Summary:

IBM is a massive company that is well versed in almost every type of AI application. They have currently created a branch that strictly focuses on AI and ML in medical diagnosis and research. One area of interest for the otologists includes IBM's work in Process Acceleration in HealthCare. This software focuses on treatment options based on symptoms listed by doctors (or patients). We believe that the technology applied here can assist the otologists in diagnosing tumors just from a set of symptoms. The solution is so advanced that it is currently undergoing medical certification and product. We also believe that IBM is an excellent resource for consultation on questions during the implementation of AI systems.

Notable Publications:

IBM has been involved in research projects from a variety of medical domains. There are currently no research projects strictly related to otological ailments. However, there has been one application of their product in predictive analysis in prostate cancer. IBM created software that takes a series of images from the cells in the prostate region and determines if the region is at a high risk for developing cancer. The product identified 5 regions that were previously known to be high risk. Additionally, it also discovered an additional 9 regions using its algorithms though logistic regression.

Contact Information:

Website: https://www.zurich.ibm.com/
Address: IBM Research GmbH Säumerstrasse 4 CH–8803 Rüschlikon, Switzerland
Phone Number: +(41) 44 724 8111

A.2.5 Bolzano

Summary:

Balzano is a company that focuses on developing AI for augmented diagnosis with medical MRIs. They specialize in classification, quantification and predictive analysis for clinical 3D images (MRI, CT), EHR and clinical documentation through deep learning or traditional machine learning technologies. They are currently working with Swiss hospitals to create a catalog of detectable conditions that are both frequent and rare. The technologies used from Balzano have been used for real-time second opinions, peer review, and case comparison. The practical use of Balzano technologies is for automatic interpretation of musculoskeletal MRI images.

Notable Publications:

Balzano is currently partnered with a variety of hospitals in the Swiss area. While they do not have any publications listed under their name, they have listed some documents on their website that detail some of the applications of their work. This can be found here.

Contact Information:

- Email: info@balzano.net
- Phone: +41 0840 840 365
- Address: Zollikerstrasse 153 8008 Zürich, Switzerland
- LinkedIn: Balzano

Appendix B: Interview Protocols

The interviews our team conducted were semi-structured. This means that while we had a prepared list of interview questions, we modified this list as the interview progressed. Each interview was therefore slightly different as we asked specific probing questions pertaining to the interviewee. For example, we asked questions about the specific workings of each otologist's research projects. These questions differed from otologist to otologist. The same is true for the two AI company interviews we conducted. Many questions pertain specifically to that company and were created as the interview progressed to get more specific knowledge in a particular area

.

Before each interview we read a consent form to the participant, then asked if we could record the interview. The consent form can be found below. After our introductions, we began asking questions. The structured questions that were asked in all interviews can be found below, organized by type of interview. We have excluded pleasantries and introductions from the interview questions.

B.1 Consent to be Interviewed

If you choose to consent to the interview, we expect this process to take approximately 60 minutes to complete. The information from this interview will be published in 2 two places:

1. Worcester Polytechnic Institute, as part of our final research report.
2. Universitätsspital Zürich, as part of our deliverable package for the Zürich Otology Department.

Your identity will not be published in our findings. The information collected from this interview will be used to inform our research into how artificial intelligence could best be implemented into the work in the University Hospital Zürich's otology department. Please note that this process is voluntary as you can refrain from participating in the interview as a whole or refrain from answering a certain question. If you wish to contact our group or university for any reason, it is as follows:

Our team can be contacted using our group email alias at gr-AISwissA20@wpi.edu.

Our university's internal review board can be contacted using the following information:

IRB Manager (Ruth McKeogh, Tel. 508-831-6699, Email: irb@wpi.edu) and the Human Protection Administrator (Gabriel Johnson, Tel. 508-831-4989, Email: gjohnson@wpi.edu)

B.2 Interview Questions for Otologists at the University Hospital Zürich

1. What projects are you currently working on?
2. What are some common issues that emerge in your work?
3. We would like to discuss the data collection involved in your work.
    a. What kinds of data are collected in your research?
    b. How do you typically collect this data?
    c. How do you typically organize and present this data?
4. Have you used artificial intelligence algorithms when conducting research?
    a. Probes:
        i. If yes:
            1. How has artificial intelligence impacted your research?
        ii. If no:
            1. Would you want to?
                a. Yes: How can you envision it helping?
                b. No: Why not?
5. Typically, AI is good at image and pattern recognition. If you had an AI assisting in your work, what would you want it to do?
6. How do you think AI could help in other aspects of otology?
7. Do you know of any products/companies that you think might be of assistance to your research?
8. What future projects are you looking to work on?

B.3 Interview Questions for AI developers/companies

1. Tell us a little bit about yourself, and [insert company name].
2. What types of artificial intelligence-based solutions does [insert company name] provide?
3. Has [insert company name] ever provided software-based solutions for organizations such as hospitals or medical research facilities?
        i. Probes:
            1. If yes:
                a. If so, whom?
                b. What types of problems were these groups looking to solve?

    2. If no:
        a. Is this an area of interest for [insert company here], and if so, what software-based tools might the company be looking to develop?

4. One otologist from UHZ is looking for an AI that can analyze data about how the head changes and deforms with variances in sound frequencies from an ear device. They take three-dimensional data on the velocities of various points on the head, then overlay that with the frequency at that time. They are looking for an artificial intelligence that can take this data and recognize the patterns and trends.

5. One area of concern includes understanding how the medical AI arrives at a particular answer/diagnosis, in other words, the machine's “explainability”. Does the AI address this issue?
    i. Probe:
        1. If yes:
            a. How so?
        2. If no:
            a. Is this an area of interest for [insert company name]?
6. Does [insert company name] have any competitors in the local Zürich area?
    i. Probes:
        1. If so, what sets [insert company name] apart from the rest of the competition?
7. If the UHZ Otology department requests your software, how might they go about implementing it?
    i. Probes:
    ii. Is it a time-intensive process?
        1. What is the cost associated with this purchase?
        2. Does it require knowledge of any external items unrelated to the field of medicine? For example, skills in any of the following: computer systems, AI networks, or software development?
8. How might future legislation regarding privacy of data affect the company’s future aspirations or products?
9. How does the software handle personal data? In what way is it protected?
10. How do you see [insert company name] influencing medical-based AI in the next 5-10 years?

Appendix C: Interview Notes

C.1 Interview with Dr. Christof Röösli on September 14th at 10:00 AM EST (4:00 PM CEST)

Current Projects:
- Clinician
- Part time research
- Bone conduction in hearing
- Material properties of skull bone
- Propagation of hearing
- Analysis of surgical outcomes and results
    - Hearing outcomes
    - How symptoms change (questionnaires)
- Classification systems

Kind of surgery
- Repair of membranes
- Repair of ossicles in middle ear
- Cochlear implant
- Electrode placed in middle ear
- Tumor removals

Common issues:
- Research
    - So much data
    - Long time to process
    - Figure out which are important

How do you currently analyze data?
- MATLAB
- Excel sheets
- Statistics

Data:
- Measure on cadavers

AI

- Have you used it before?
    - Doesn't know much
    - Wants to plan a project
    - Needs more experience
    - Read articles
- What do you want it to do?
    - Predict if a tumor would grow given an image
    - How likely it would be to grow
        - Saves time and money (years and many MRIs)
    - MRI tracks the tumor growth
    - Record diameter and volume
    - Outcome of surgery given multiple factors
        - Needs multiple data points (unsure of what kinds of data)
    - Other areas of otology?
        - Tumors anywhere
        - Image analysis
    - Tool that could calculate with the data (big data)

Future projects:
- Data analysis
- Bone conduction
- Clinical application

C.2 Interview with Dr. Christof Röösli on September 28th at 10:00 AM EST (4:00 PM CEST)

Tumor research:
- How often do you take the MRIs?
    - 50 new cases a year
    - Typically repeat MRI 6 months later and check if it grows, if not then 1 MRI a year, then every 2 years, then every 5 as if still doesn't grow
    - Keep these images
    - Search system for specific patients
    - Doesn't separate by if they are growing or not
    - Primary focus: vestibular schwannoma
    - Found in canal between ear and brain
- Tools
    - Measure diameter

    - Seeing structure is not too important
    - Cystic tumors, homogeneous
        - Can only really see darker and lighter spots
    - Fluid would be visible
    - Usually some solid and some cystic parts
    - Fluid shows as brighter
- Images
    - Spatial resolution and size?
        - Email that question
- Training AI
    - Many images already stored
    - 100 new scans a year approximately
    - Many years, a lot of data, image quality has changed a bit
        - Look up exactly how many new scans (email question)
    - Other otologists likely also keep these scans and information
        - Images are accessible
            - Ethical approval required
    - Key things to pick up for AI
        - If it's growing or not is the big problem / question
            - Only way to tell is comparing scans
- Other tests being done?
    - Hearing tests
        - Audiograms and speech
    - Vestibular test
    - Discrepancy here prompts MRI
    - Hearing loss does not correlate to tumors / size
- Also do surgery to remove tumors
    - Craniotomy
    - Go through the ear
        - Vestibular
    - Tumor needs to be mobilized from facial nerve
    - 6-hour surgery typically