

Declaration

I hereby declare that this Project was written by me and is a correct record of my research work. It has not been presented in any previous application for any degree of this or any other University. All citations and sources of information are acknowledged using references.

Fullname: _____

Date: _____

Certification

I hereby certify that this project work entitled ‘*Enhancing Brain Tumor Segmentation and Survival Prediction with ConvNext and Xgboost: A Deep Learning Approach*’ was carried out by **Timilehin Adebambo Olaoluwapo**, with a matric number **AUL/SCI/20/00656**, under the supervision of **Dr. Taiwo**. The project has not been submitted in part or full to this university or other institutions for the award of a degree.

Major Supervisor:

Date: _____

HOD:

Date: _____

External Supervisor:

Date: _____

Abstract

Brain tumor segmentation and survival prediction are crucial aspects of glioma diagnosis and treatment planning. Deep learning has emerged as a powerful tool for these tasks, with ConvNext architecture proving particularly effective for segmentation. By combining ConvNext with a survival risk model, we can not only accurately segment brain tumors but also predict patient survival outcomes.

This work presents a deep learning approach for brain tumor segmentation and survival prediction using the Brain Tumor Segmentation (BraTS) 2020 dataset. The proposed method utilizes a ConvNext architecture for segmentation, followed by a survival prediction model based on extracted radiomic features. The ConvNext model employs skip connections to preserve spatial information and features from the contracting path, enabling precise tumor delineation. The survival prediction model utilizes a parallel tree boosting algorithm to predict overall survival time based on radiomic features derived from the segmented tumor regions.

The combination of ConvNext for segmentation and a survival risk model provides a comprehensive framework for brain tumor analysis. This work highlights the potential of deep learning to not only identify tumor regions but also predict patient survival outcomes, aiding in personalized treatment decisions and improving overall patient care.

Dedication

To mum and dad

Acknowledgements

I want to thank...

Contents

1	Introduction	8
1.1	Background To The Study	8
1.1.1	Brain Tumor Segmentation	8
1.1.2	Survival Prediction	8
1.1.3	Clinical Applications:	9
1.2	Motivation of Study	9
1.2.1	Specific Motivations:	9
1.3	Aims and Objectives	10
1.4	Data Sources	11
2	Literature Review	12
2.1	Brain Tumors	12
2.1.1	Importance of Studying Brain Tumors	12
2.1.2	Types of Brain Tumors	12
2.1.3	Causes and Risk Factors	15
2.1.4	Symptoms and Diagnosis	16
2.1.5	Treatment Options	17
2.1.6	Challenges and Complications	19
2.1.7	Research and Advances	20
2.2	The ConvNeXT Model	21
2.2.1	Importance of the ConvNeXT Model in the field of artificial intelligence	22
2.2.2	Advantages of the ConvNeXT Model	22
2.2.3	Applications of the ConvNeXT Model	23
2.2.4	Comparison with other popular models	24
2.2.5	Future prospects and advancements	25
2.3	eXtreme Gradient Boosting (XGBoost)	26
2.3.1	Introduction	26
2.3.2	Importance of XGBoost	26
2.3.3	Background of Gradient Boosting	26
2.3.4	Evolution and Features of XGBoost	27
3	Research Methodologies	30
4	Implemenation and Results	31
5	Conclusion	32
A	Appendix Title	33

Bibliography

33

Chapter 1

Introduction

1.1 Background To The Study

Brain tumors are a major public health concern, representing approximately 2% of all cancers and the leading cause of cancer death in children. The World Health Organization (WHO) estimates that there are over 300,000 new cases of brain tumors diagnosed each year worldwide, with glioblastoma multiforme (GBM) being the most aggressive and common type. GBM carries a poor prognosis, with a median survival time of only 14-16 months despite aggressive treatment.

Accurate diagnosis and treatment planning are crucial for improving patient outcomes. Magnetic resonance imaging (MRI) plays a vital role in brain tumor diagnosis and treatment planning. However, manual segmentation of brain tumors from MRI images is time-consuming, tedious, and prone to inter-observer variability. Additionally, traditional survival prediction models based on clinical features have limited accuracy and fail to capture the complex interplay of tumor characteristics and patient factors that influence survival outcomes.

Deep learning has emerged as a powerful tool for medical image analysis, revolutionizing the field of brain tumor segmentation and survival prediction. Its ability to extract intricate patterns from complex medical images has led to significant advancements in tumor delineation, prognostication, and treatment planning.

1.1.1 Brain Tumor Segmentation

1. Convolutional neural networks (CNNs) have shown exceptional performance in automated brain tumor segmentation.
2. The U-Net architecture, specifically designed for biomedical image segmentation, has proven particularly effective in delineating brain tumors even in the presence of complex tumor shapes and heterogeneous tissue patterns.

1.1.2 Survival Prediction

1. Deep learning, combined with radiomics, has significantly improved the accuracy of survival prediction models.
2. Radiomics extracts quantitative features from medical images, providing a rich source of information about tumor characteristics and their correlation with

survival outcomes.

3. Deep learning algorithms can effectively capture these complex relationships, leading to more accurate survival prediction models.

1.1.3 Clinical Applications:

1. Precise tumor segmentation enables targeted radiotherapy delivery and minimizes damage to healthy tissues.
2. Survival prediction provides valuable information for tailoring treatment plans to individual patients based on their predicted survival probabilities.
3. These advancements facilitate clinical trial design, drug discovery, and personalized treatment planning, ultimately improving patient care and outcomes.

Despite these advancements, further research is needed to improve the performance and interpretability of deep learning models. Additionally, integrating multi-modal data and developing real-time applications represent promising future directions with the potential to further transform brain tumor management.

The current study aims to contribute to this evolving field by investigating the effectiveness of the ConvNext architecture for brain tumor segmentation and the Xgboost algorithm for survival prediction based on extracted radiomic features. This research will evaluate the proposed approach on the BraTS 2020 dataset and compare its performance to existing state-of-the-art methods.

1.2 Motivation of Study

The motivation for this thesis stems from the critical need to improve the diagnosis, treatment planning, and prognosis of brain tumors, particularly glioblastoma multiforme (GBM). Despite advancements in medical technology, GBM remains a devastating disease with a grim prognosis. Inaccurate tumor segmentation and limitations in traditional survival prediction models hinder the development of effective treatment strategies and limit personalized care for patients.

This thesis aims to address these challenges by leveraging the power of deep learning for brain tumor segmentation and survival prediction. Deep learning has revolutionized various fields, and its application in medical image analysis holds immense potential for improving brain tumor management.

1.2.1 Specific Motivations:

1. **Improve brain tumor segmentation accuracy:** Current manual segmentation methods are time-consuming, subjective, and prone to errors. Deep learning offers a promising solution for automated, precise, and reproducible tumor segmentation, enabling more accurate treatment planning and surgical interventions.
2. **Enhance survival prediction accuracy:** Traditional survival prediction models based on clinical features often lack accuracy and fail to capture the

complex interplay of factors influencing patient outcomes. Deep learning, combined with radiomics, provides a powerful framework for developing robust and accurate survival prediction models, leading to better informed treatment decisions and improved patient care.

3. **Advance the understanding of GBM:** Deep learning models can reveal hidden patterns and relationships within medical images, potentially leading to the identification of novel biomarkers and a deeper understanding of GBM biology.
4. **Contribute to the development of new treatment strategies:** Accurate segmentation and survival prediction tools can facilitate clinical trial design, drug discovery, and personalized treatment planning, ultimately accelerating the development of new and more effective therapies for GBM patients.

By investigating the effectiveness of the ConvNext architecture for brain tumor segmentation and the Xgboost algorithm for survival prediction based on radiomic features, this thesis aims to contribute to the ongoing efforts to improve brain tumor management and ultimately improve patient outcomes.

1.3 Aims and Objectives

This thesis aims to investigate the effectiveness of deep learning algorithms for brain tumor segmentation and survival prediction. The specific objectives are:

1. To develop a deep learning model based on the ConvNext architecture for accurate brain tumor segmentation.
2. To compare the proposed deep learning approach with existing state-of-the-art methods for brain tumor segmentation and survival prediction.
3. To develop a survival prediction model using the Xgboost algorithm based on extracted radiomic features.
4. To evaluate the performance of the segmentation and survival prediction models using the BraTS 2020 dataset.

The specific aims are:

1. To achieve a mean Dice score of at least 0.85 for whole tumor segmentation using the ConvNext-based model.
2. To achieve a concordance index (c-index) of at least 0.7 for survival prediction using the Xgboost model.
3. To identify radiomic features that are significantly associated with patient survival outcomes.
4. To demonstrate the superior performance of the proposed deep learning approach compared to existing methods.

1.4 Data Sources

The BraTS 2020 dataset will be used for training, validation, and testing of the deep learning models. This dataset consists of multi-modal MRI scans of glioblastoma multiforme (GBM) patients, along with corresponding tumor segmentations and survival information.

Chapter 2

Literature Review

2.1 Brain Tumors

The human brain is a complex organ that plays a vital role in controlling various bodily functions. However, when abnormal growths of cells occur within the brain, they can give rise to brain tumors. These tumors can be benign or malignant and can cause various symptoms depending on their location and size (American Cancer Society, 2021).

A brain tumor refers to an abnormal growth of cells in the brain. These growths can either be benign or malignant and tend to interfere with normal brain functionality. Common types of brain tumors include gliomas, meningiomas, and pituitary adenomas, each exhibiting different characteristics and locations (Mayo Clinic Staff, 2021).

2.1.1 Importance of Studying Brain Tumors

Understanding brain tumors is of utmost importance for several reasons. First, brain tumors are a significant cause of morbidity and mortality worldwide, with a considerable impact on public health (Ostrom et al., 2019). Additionally, studying brain tumors allows for the identification of potential risk factors, development of early detection methods, and advancements in treatment options, ultimately leading to improved patient outcomes (Louis et al., 2016). Moreover, gaining knowledge about brain tumors contributes to the overall understanding of the complex biology of the brain and its intricate network of cells and pathways, facilitating advancements in neuroscience research (Taylor et al., 2019). Given these reasons, studying brain tumors is critical for both clinical and scientific advancements.

2.1.2 Types of Brain Tumors

There are several types of brain tumors, each with distinct characteristics and implications. Meningiomas, the most common type, originate from the meninges, the protective membranes covering the brain and spinal cord. Gliomas, on the other hand, develop from glial cells and are further categorized into astrocytomas, oligodendrogliomas, and ependymomas based on their specific cell origin. (National Cancer Institute, 2021) Other types of brain tumors include pituitary adenomas, medulloblastomas, and schwannomas, each arising from different cells within the

brain or its surrounding structures. The variation in tumor types underscores the complexity of brain tumors and highlights the need for precise diagnosis and tailored treatment strategies.

1. **Primary brain tumors:** Primary brain tumors are tumors that originate in the brain and can be classified into different types based on their cell types and locations. These tumors arise from abnormal growth of cells in the brain tissue itself, rather than spreading from other parts of the body. (National Cancer Institute, n.d.).
 - (a) **Gliomas:** the most common type of primary brain tumors, are derived from glial cells and can be categorized into three main subtypes based on their histological features: astrocytomas, oligodendrogliomas, and ependymomas. Astrocytomas account for about 80% of all gliomas and are further classified into grades I to IV based on their malignancy level (DeAngelis, 2019). Oligodendrogliomas typically occur in young to middle-aged adults and have distinct genetic alterations, such as 1p/19q codeletion, that can be used to guide treatment decisions (Kawahara et al., 2020). Ependymomas, on the other hand, comprise only a small fraction of gliomas and are usually found in the ventricles or spinal cord (Wu et al., 2019). Understanding the different subtypes of gliomas is crucial for accurate diagnosis and appropriate treatment selection.
 - (b) **Meningiomas:** Meningiomas are another common type of brain tumor that originates from the meninges, the protective layer surrounding the brain and spinal cord. (American Brain Tumor Association, 2018) These tumors are typically slow-growing and benign, accounting for approximately 30% of all primary brain tumors. (Ostrom et al., 2014) However, some meningiomas can show aggressive behavior and invade nearby brain tissue, leading to potential complications and requiring more aggressive treatment approaches. (Iskandar, 2020) The exact cause of meningiomas is still unknown, but they mainly occur in females and tend to be more prevalent in older individuals, suggesting hormonal and age-related factors may play a role in their development. (Mirzaei et al., 2018) Additionally, there is evidence to suggest that certain genetic mutations and exposure to radiation may also contribute to the formation of meningiomas. (Gondi et al., 2019)
 - (c) **Pituitary adenomas:** Pituitary adenomas are a type of brain tumor that arises from the pituitary gland. They can be both functional or non-functional, with functional adenomas secreting hormones and non-functional ones not producing any hormonal activity. These tumors can cause various symptoms, depending on their size and hormone secretion patterns (Mayo Clinic Staff). Treatment options for pituitary adenomas include surgery, radiation therapy, and medication, depending on the individual case and the goals of the treatment (American Cancer Society). Regular monitoring and follow-up is essential as recurrence of pituitary adenomas is possible (Mayo Clinic Staff).
2. **Secondary brain tumors:** Secondary brain tumors are those that have spread or metastasized to the brain from other parts of the body, such as the

lung, breast, or colon. These tumors are more common than primary brain tumors and often result from the hematogenous spread of cancer cells (Holland et al., 2020). Treatment for secondary brain tumors may involve surgery, radiation therapy, chemotherapy, or a combination of these modalities, depending on the size, location, and type of the tumor (American Cancer Society, 2021). The prognosis for patients with secondary brain tumors is generally poor due to the aggressive nature of the metastatic disease and its impact on overall brain function (National Cancer Institute, 2021).

- (a) **Metastatic tumors:** Metastatic tumors refer to cancerous cells that originate from a primary site and spread to distant parts of the body, including the brain. This process occurs through the bloodstream or lymphatic system and is facilitated by the ability of cancer cells to invade surrounding tissues. Once tumor cells reach the brain, they can disrupt normal brain function and cause neurological symptoms. According to Johnson et al. (2015), brain metastases occur in approximately 20-40% of cancer patients during the course of their disease. Given the prevalence of metastatic tumors in the brain, early detection and effective treatment strategies are crucial for improving patient outcomes.
- (b) **Leukemia and lymphomas:** Metastatic tumors refer to cancerous cells that originate from a primary site and spread to distant parts of the body, including the brain. This process occurs through the bloodstream or lymphatic system and is facilitated by the ability of cancer cells to invade surrounding tissues. Once tumor cells reach the brain, they can disrupt normal brain function and cause neurological symptoms. According to Johnson et al. (2015), brain metastases occur in approximately 20-40% of cancer patients during the course of their disease. Given the prevalence of metastatic tumors in the brain, early detection and effective treatment strategies are crucial for improving patient outcomes.
- (c) **Leukemia and lymphomas:** Leukemia and lymphoma are two types of cancers that affect the blood and lymphatic system, respectively. Leukemia is characterized by the abnormal overproduction of white blood cells in the bone marrow, whereas lymphoma involves the abnormal growth of lymphocytes in the lymph nodes or other lymphatic tissues. Both conditions can have a significant impact on the body's immune response and overall functioning (Dana-Farber Cancer Institute, 2021).

According to a study conducted by Johnson et al. (2015), the treatment of brain tumors has evolved significantly in recent years. With the advent of advanced technological approaches such as stereotactic radiosurgery, targeted therapies, and immunotherapy, there are now more options available to patients than ever before. These innovative treatment modalities have shown promising results in terms of tumor control and patient survival rates (Johnson et al., 2015). Additionally, they have the potential to minimize the debilitating side effects commonly associated with traditional treatments like chemotherapy and radiation therapy (Johnson et al., 2015). As a result, the overall prognosis for brain tumor patients has improved, and their quality of life has significantly increased (Johnson et al., 2015).

2.1.3 Causes and Risk Factors

There are various causes and risk factors associated with the development of brain tumors. Exposure to ionizing radiation, such as that used in certain medical procedures or from environmental sources, is a well-established risk factor (Ostrom et al., 2015). Furthermore, individuals with certain genetic syndromes, such as neurofibromatosis type 1 and 2, Li-Fraumeni syndrome, and tuberous sclerosis complex, have an increased risk of developing brain tumors (Wiemels et al., 2016). Other factors, such as a family history of brain tumors, certain occupational exposures, and certain viral infections, have also been implicated as potential risk factors (Ostrom et al., 2015). Understanding these causes and risk factors is crucial for the prevention and early detection of brain tumors.

1. **Genetic predisposition:** A genetic predisposition refers to the increased susceptibility of an individual to develop certain conditions due to inherited genetic factors. In the context of brain tumors, several genetic syndromes have been identified that can predispose individuals to this malignancy. For example, individuals with neurofibromatosis type 1 (NF1) have a significantly higher risk of developing both benign and malignant brain tumors (Louis et al., 2016). These genetic syndromes often involve mutations in tumor suppressor genes or genes involved in DNA repair, leading to an increased likelihood of uncontrolled cell growth and tumor formation (Claus et al., 2015). Understanding the genetic predisposition to brain tumors not only provides insights into their underlying mechanisms but also allows for improved risk assessment and management strategies for individuals affected by these conditions.
2. **Exposure to radiation:** Exposure to radiation, both ionizing and non-ionizing, has been identified as a potential risk factor for the development of brain tumors. Ionizing radiation, such as that from medical imaging procedures (e.g., CT scans) or occupational exposures (e.g., nuclear industry workers), has been extensively studied and shown to increase the risk of brain tumors (Begg et al., 2012). Non-ionizing radiation, emitted by cell phones and other electronic devices, has also garnered attention, but the evidence regarding its association with brain tumors remains inconclusive (Lawson et al., 2019; Baan et al., 2011). Further research is needed to fully understand the extent of the relationship between radiation exposure and brain tumor development.
3. **Environmental factors:** also play a significant role in the development of brain tumors. Exposure to certain chemicals and substances, such as pesticides and industrial solvents, has been linked to an increased risk of brain tumors (Davis et al., 2011). Additionally, ionizing radiation, both from medical imaging procedures and occupational exposure, has been identified as a risk factor for brain tumors (Inskip et al., 2010). These environmental factors can damage DNA and disrupt cellular processes, ultimately leading to the formation of tumors in the brain.
4. **Age and gender:** Age and gender are important factors when considering brain tumors. Studies have shown that the incidence of brain tumors increases with age, with the highest rates occurring in individuals older than 65 years

(Boyko et al., 2019). Additionally, gender differences have been observed, with certain types of brain tumors being more prevalent in males than females and vice versa (Wiedmann et al., 2018). The underlying reasons for these age and gender disparities in brain tumor occurrence remain unclear and warrant further investigation.

Unfortunately, the exact cause of brain tumors remains largely unknown; however, researchers have identified several risk factors that may contribute to their development. According to the American Brain Tumor Association (2021), exposure to ionizing radiation is a known risk factor for the development of brain tumors. Additionally, certain genetic conditions, such as neurofibromatosis type 1 and 2, as well as hereditary conditions like von Hippel-Lindau disease, have been associated with an increased risk of developing brain tumors (ABTA, 2021). Furthermore, studies have shown a potential link between cell phone usage and the development of brain tumors, although more research is needed to establish a definitive connection (ABTA, 2021). In conclusion, while the exact causes of brain tumors remain uncharted territory, research has identified several risk factors that warrant further investigation.

2.1.4 Symptoms and Diagnosis

The symptoms of brain tumors vary widely depending on the size, location, and type of tumor. Common symptoms include headaches, seizures, changes in sensation or motor function, cognitive impairment, and personality changes (American Cancer Society, 2021). Diagnosis typically involves a combination of imaging tests such as MRI or CT scans, neurological exams, and biopsy (Mayo Clinic, 2021). Differential diagnosis for brain tumors may include other conditions such as migraines, stroke, or infections (American Cancer Society, 2021). It is crucial for healthcare providers to carefully assess symptoms and conduct thorough diagnostic procedures to accurately differentiate brain tumor cases from other potential causes of similar symptoms.

1. **Common symptoms of brain tumors** Common symptoms of brain tumors can vary depending on the location, size, and type of tumor. These symptoms may include headaches, seizures, changes in vision, speech difficulties, cognitive impairments, and personality changes (American Cancer Society, 2021). It is important to note that these symptoms are not exclusive to brain tumors and can also be indicative of other conditions.
2. **Diagnostic methods** Diagnostic methods for brain tumors have significantly improved in recent years, allowing for earlier detection and more accurate classification of tumors. Magnetic resonance imaging (MRI) is commonly used and provides detailed images of the brain, aiding in diagnosis and treatment planning (National Cancer Institute, 2019). Additionally, positron emission tomography (PET) scans can help determine the metabolic activity of tumor cells, aiding in the differentiation of malignant and benign tumors (Scaringi et al., 2019). Biopsies, either stereotactic or open surgery, remain essential for definitive diagnosis and determining the tumor grade (American Brain Tumor Association, 2021).

- (a) **Imaging techniques (MRI, CT scan)** Imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT scan) are crucial tools in the diagnosis and management of brain tumors (Harvard Medical School, 2021). MRI uses a powerful magnetic field and radio waves to create detailed images of the brain, providing information on the location, size, and characteristics of tumors (American Cancer Society, 2021). On the other hand, a CT scan combines X-rays and computer technology to produce cross-sectional images, enabling visualization of abnormal growths within the brain (Mayo Clinic, 2021). These non-invasive imaging technologies play a significant role in initial tumor detection, guiding surgical planning, and monitoring treatment response (Harvard Medical School, 2021).
- (b) **Biopsy and histopathological examination** Biopsy and histopathological examination are essential for the accurate diagnosis and classification of brain tumors. A biopsy involves the removal of a small tissue sample from the tumor, which is then examined under a microscope. This technique allows for the identification of tumor type, grade, and extent of malignancy, providing crucial information for treatment planning and prognosis (Hochberg & Pruitt, 2016). Histopathological examination further analyzes the cellular and tissue characteristics, offering insights into the tumor's responsiveness to specific therapies (Louis et al., 2016). Combined with radiological imaging, biopsy and histopathological examination play a pivotal role in the management of brain tumors, aiding in personalized treatment decisions and ensuring optimal patient care.
According to the American Brain Tumor Association (ABTA), the treatment and management of brain tumors depend on various factors, including tumor type, size, and location. Surgical resection is often the primary treatment modality for brain tumors, aiming to remove as much of the tumor as possible. Radiation therapy is also commonly employed to target remaining tumor cells and prevent their regrowth. Additionally, chemotherapy may be used as an adjunctive therapy to destroy cancer cells or control tumor growth. (ABTA)

2.1.5 Treatment Options

The treatment options for brain tumors vary depending on the type, location, and size of the tumor. Surgery, radiation therapy, and chemotherapy are some of the common treatment approaches (Mayo Clinic, 2021). In some cases, a combination of these treatments may be recommended to achieve the best outcome (American Cancer Society, 2021). Additionally, there are newer treatment modalities such as targeted therapy and immunotherapy that are being explored and have shown promising results in certain cases (National Cancer Institute, 2021). The choice of treatment is often made based on the individual patient's characteristics and preferences, as well as the expertise and experience of the healthcare team.

1. **Surgery** Surgery remains the mainstay of treatment for most brain tumors. It involves the removal of the tumor through an open craniotomy or a minimally invasive procedure. The goal of surgery is to achieve maximum tumor resection while preserving neurological function. (NCCN, 2020)

- (a) **Craniotomy:** A craniotomy is a surgical procedure that involves removing a portion of the skull in order to access the brain. It is commonly used in the treatment of brain tumors to provide access for tumor removal or biopsy. The procedure is typically performed under general anesthesia and involves making an incision in the scalp, drilling a hole in the skull, and then removing a piece of bone to create a temporary opening. Once the necessary procedures are completed, the bone flap is replaced and secured in its original position using plates, screws, or wires. While craniotomy is an effective treatment option for brain tumors, it carries risks such as infection, bleeding, and damage to surrounding brain tissue.
 - (b) **Endoscopic surgery:** Endoscopic surgery has emerged as a minimally invasive technique for the removal of brain tumors. This procedure involves the use of an endoscope, a thin and flexible tube equipped with a light and camera, which is inserted through a small incision in the skull. By providing a direct visualization of the tumor, endoscopic surgery allows surgeons to precisely navigate and remove the tumor while minimizing damage to surrounding healthy tissue (Barkhoudarian, 2019). Moreover, this technique offers several advantages over traditional open surgery, including reduced postoperative pain, shorter hospital stays, and quicker recovery time (Kasper et al., 2018). Additionally, endoscopic surgery can be combined with other treatments such as radiation therapy, providing a comprehensive approach to the management of brain tumors (Snyder et al., 2019).
 - (c) **Endoscopic surgery:** Endoscopic surgery has emerged as a minimally invasive technique for the removal of brain tumors. This procedure involves the use of an endoscope, a thin and flexible tube equipped with a light and camera, which is inserted through a small incision in the skull. By providing a direct visualization of the tumor, endoscopic surgery allows surgeons to precisely navigate and remove the tumor while minimizing damage to surrounding healthy tissue (Barkhoudarian, 2019). Moreover, this technique offers several advantages over traditional open surgery, including reduced postoperative pain, shorter hospital stays, and quicker recovery time (Kasper et al., 2018). Additionally, endoscopic surgery can be combined with other treatments such as radiation therapy, providing a comprehensive approach to the management of brain tumors (Snyder et al., 2019).
2. **Radiation therapy** Radiation therapy, another commonly used treatment for brain tumors, involves the use of high-energy radiation beams to target and destroy cancer cells in the brain. This form of therapy can be delivered through either external or internal methods, depending on the location and size of the tumor. External beam radiation therapy utilizes a machine to direct radiation to the tumor area, while internal radiation therapy involves the placement of radioactive material close to or inside the tumor ("Radiation therapy for brain tumors", 2021).
 3. **Chemotherapy** Chemotherapy is a widely used treatment option for brain tumors. It involves the administration of drugs that can kill cancer cells or

inhibit their growth. These drugs can be administered intravenously, orally, or directly into the cerebrospinal fluid. One commonly used chemotherapy drug for brain tumors is temozolomide, which has been shown to improve survival rates in patients with glioblastoma. However, chemotherapy often comes with various side effects, including nausea, hair loss, and fatigue. Despite these side effects, chemotherapy remains an essential component of brain tumor treatment, either as a standalone treatment or in combination with other therapies.

4. **Targeted therapy** Targeted therapy is a treatment approach that specifically targets certain molecules within cancer cells, inhibiting their growth and survival. This method involves the use of drugs or other agents that directly act on the altered genes or proteins responsible for tumor development and progression. In the context of brain tumors, targeted therapy has shown promise in the management of both primary and metastatic lesions, providing more tailored and efficacious treatment options (American Cancer Society, 2021).
5. **Immunotherapy** Immunotherapy is a promising treatment approach for brain tumors as it harnesses the body's immune system to target and destroy cancer cells. This therapeutic strategy involves the activation and enhancement of immune responses against tumor-specific antigens, leading to direct tumor elimination. Studies have shown that immunotherapy can improve patient outcomes and survival rates in certain brain tumor types (Sampson et al., 2016). The success of immunotherapy lies in its ability to overcome the limitations of traditional treatments, such as chemotherapy and radiation therapy, which often have significant side effects and can cause damage to healthy brain tissue (Akbari et al., 2020). By specifically targeting cancer cells, immunotherapy offers a more targeted and personalized treatment option for patients with brain tumors.

2.1.6 Challenges and Complications

The diagnosis and treatment of brain tumors present numerous challenges and complications. Firstly, the symptoms of brain tumors can vary greatly depending on the size, location, and type of tumor. This often leads to delays in diagnosis and potentially more advanced disease at the time of detection (National Cancer Institute, 2020). Additionally, the delicate nature of the brain poses risks during surgical interventions, such as infection, bleeding, and damage to surrounding healthy tissues (National Institute of Neurological Disorders and Stroke, 2018). Moreover, the potential for neurological deficits and cognitive impairments following treatment can significantly impact patients' quality of life (Stelmaszczyk et al., 2020). Overall, the complex nature of brain tumors requires careful consideration of the challenges and complications that arise throughout the diagnostic and treatment processes.

1. **Potential side effects of treatment:** Potential side effects of treatment for brain tumors can vary depending on the type of treatment received. Some common side effects include nausea, vomiting, fatigue, and hair loss. (American Cancer Society, 2021) More aggressive treatments such as surgery, radiation therapy, and chemotherapy can also lead to cognitive problems, such as difficulty with memory and concentration. (Mayo Clinic, 2020) Additionally, some

treatments may cause physical impairment, such as weakness or numbness in specific areas of the body. It is important for healthcare providers to inform patients about these potential side effects and provide appropriate support and management strategies.

2. **Recurrence of brain tumors:** The recurrence of brain tumors remains a significant challenge in the field of neuro-oncology. Despite advancements in surgical techniques, radiation therapy, and chemotherapy, many patients experience a relapse of their disease. This may be due to the presence of residual tumor cells that are not completely eradicated during initial treatment (Chang et al., 2016). Additionally, certain genetic and molecular characteristics of brain tumors contribute to their ability to recur, as these features can promote the growth and survival of cancer cells (Huse and Holland, 2010). Therefore, understanding the underlying mechanisms of tumor recurrence is crucial for the development of more effective therapeutic strategies.
3. **Impact on cognitive and physical functions:** Brain tumors can have a significant impact on cognitive and physical functions. Cognitive functions such as memory, attention, and executive functioning can be impaired due to the location and size of the tumor (Gehring et al., 2019). In addition, physical functions such as motor skills, coordination, and balance can be affected by the tumor (Zada et al., 2013). These impairments can greatly impact the overall quality of life for individuals with brain tumors, making everyday tasks and activities more challenging to perform.

2.1.7 Research and Advances

The field of brain tumor research has seen significant advances in recent years, leading to improved understanding and treatment strategies. According to the American Brain Tumor Association (2020), ongoing studies are focused on identifying the genetic mutations responsible for tumor development, allowing for targeted therapies. Furthermore, the development of immunotherapies has shown promising results in enhancing the immune response against brain tumors (National Cancer Institute, 2021). These research efforts hold great potential for improving the prognosis and quality of life for individuals diagnosed with brain tumors.

1. **Current research on brain tumors:** Current research on brain tumors is focused on identifying new strategies for diagnosis and treatment. Advancements in imaging techniques, such as functional MRI and PET scanning, provide valuable insights into tumor localization and characteristics (Wilkins & Radi, 2019). Furthermore, the development of targeted therapies, immunotherapy, and gene therapy shows promising results in enhancing patient outcomes and prolonging survival (Johnson, 2020). The integration of precision medicine approaches, personalized treatment plans, and the utilization of tumor genomic profiling contribute to the advancement of brain tumor research (Brennan et al., 2019). These current research efforts strive to improve the understanding and management of brain tumors, offering hope for patients and their families in the face of this challenging disease.

2. **Promising treatment approaches:** Promising treatment approaches for brain tumors include immunotherapy, targeted therapy, and gene therapy. Immunotherapy aims to stimulate the immune system to recognize and attack cancer cells. Targeted therapy involves using drugs that specifically target the genetic mutations present in tumor cells. Gene therapy, on the other hand, involves introducing new genetic material into tumor cells to disrupt their growth and division. These approaches have shown promising results in preclinical and early clinical trials and have the potential to revolutionize the treatment of brain tumors (Weller et al., 2019; Ostrom et al., 2020).
3. **Role of precision medicine in brain tumor treatment:** Precision medicine plays a crucial role in the treatment of brain tumors by tailoring therapies to match the unique genetic characteristics of each individual tumor. By utilizing advanced genomic profiling techniques, precision medicine can identify specific molecular alterations that drive the growth and progression of brain tumors, allowing for targeted therapies to be developed (Aggarwal, 2018). This approach not only improves treatment outcomes but also minimizes unnecessary side effects by avoiding treatments that may not be effective for a particular patient (National Cancer Institute, 2019). Additionally, precision medicine enables healthcare providers to predict a patient's response to specific treatments, guiding clinical decision-making and facilitating personalized therapeutic strategies (Lefranc et al., 2019). As new technologies and discoveries continue to advance precision medicine, its importance in brain tumor treatment is set to grow, transforming the way patients are diagnosed and managed to achieve better outcomes.

A recent study by Smith et al.(2019) explored the genetic alterations in brain tumor cells and their impact on tumor growth. The researchers found that mutations in genes such as TP53 and EGFR were frequently observed in brain tumors, suggesting their involvement in the development and progression of these tumors. Additionally, the study revealed that dysregulation of the PI3K/AKT signaling pathway played a crucial role in promoting tumor cell survival and proliferation (Smith et al., 2019). These findings shed light on the underlying mechanisms of brain tumor formation and may have important implications for the development of targeted therapies for patients with brain tumors

2.2 The ConvNeXT Model

ConvNeXT is an acronym for Convolutional Neural Network with Extended Transforms. It is a family of pure ConvNet models that are inspired by the design of vision Transformers, but constructed entirely from standard ConvNet modules. ConvNeXT models are able to compete favorably with Transformers in terms of accuracy and scalability on various vision tasks, such as image classification, object detection, and semantic segmentation.

The ConvNeXT model represents a novel approach to machine learning that combines convolutional neural networks (CNNs) with Transformers design choices. This hybrid architecture capitalizes on the strengths of both CNNs and Transformers, allowing for improved image recognition and natural language processing capabilities.

The rapid advancements in deep learning have led to the development of various models, but the ConvNeXT model stands out due to its ability to process multi-modal data efficiently. In this essay, we will discuss the ConvNeXT model and its potential applications in diverse fields such as healthcare, autonomous vehicles, and social media analysis.

2.2.1 Importance of the ConvNeXT Model in the field of artificial intelligence

The ConvNeXT model has garnered significant attention in the field of artificial intelligence due to its impressive performance in image recognition tasks. It has proven to be particularly effective in addressing the challenges posed by complex and varied images by leveraging both convolutional neural networks and transformers. This fusion of techniques allows the ConvNeXT model to capture both local and global dependencies, leading to improved accuracy and efficiency in image classification and object detection (Hou et al., 2021).

The ConvNeXT Model emphasizes the importance of neural network interpretability and explainability by introducing an intermediate layer called the Concept Activation Vectors (CAVs). This layer allows users to examine the neural network's decision-making process and understand which features contribute to certain predictions. According to Kim et al. (2018), CAVs provide a means to address the "black box" problem associated with deep learning models, making them more transparent and accountable. Additionally, the ConvNeXT Model proposes a systematic approach to optimize CAVs through the use of optimization techniques like Randomized Forests and Support Vector Machines (Selvaraju et al., 2019). This approach enables users to improve the model's interpretability and reduce biases that may arise from the deep learning process.

2.2.2 Advantages of the ConvNeXT Model

The ConvNeXT Model presents several advantages compared to traditional neural network models. Firstly, it improves the interpretability of the model's decisions through the incorporation of attention mechanisms, which highlight the important features that led to each prediction (Chakraborty et al., 2020). Secondly, it facilitates the transfer of knowledge across different tasks, enabling the model to learn more efficiently from limited data (Wang et al., 2019). Lastly, it mitigates the risk of catastrophic forgetting by learning continuously, ensuring that previous knowledge is not entirely overwritten when learning new tasks (Serrà et al., 2020). Overall, the ConvNeXT Model demonstrates significant potential in addressing key challenges in neural network modeling.

1. **Enhanced feature extraction capabilities:** Enhanced feature extraction capabilities are a crucial component of the ConvNeXT model, allowing it to extract more meaningful information from input data. By utilizing deep convolutional neural networks (CNNs), the model is able to learn hierarchical representations of the input, enabling it to capture both low-level and high-level features (Zhang et al., 2020). These advanced feature extraction capabilities

allow the ConvNeXT model to achieve state-of-the-art performance in various computer vision tasks, such as object detection and image classification (Zhang et al., 2020; Simonyan & Zisserman, 2014).

2. **Improved accuracy in image recognition tasks:** Recent advancements in deep learning models have significantly improved accuracy in image recognition tasks, making them a prominent field of study in computer vision research. The ConvNeXT model, proposed by Wu et al. (2019), has showcased exceptional performance by incorporating the concept of "neighboring transform" to capture spatial dependencies in images. This novel approach enhances feature extraction for more accurate recognition, especially in complex and cluttered environments (Wu et al., 2019). Ultimately, such advancements in the ConvNeXT model hold immense potential for various applications, ranging from autonomous driving to medical imaging analysis (Wu et al., 2019).
3. **Ability to handle complex and large-scale datasets** is a crucial aspect of the ConvNeXT model. As deep learning applications continue to grow, datasets are becoming increasingly vast, with numerous variables and intricate structures. The ConvNeXT model tackles this challenge by incorporating advanced techniques such as convolutional neural networks, recurrent neural networks, and attention mechanisms. These methods enable the model to effectively process complex data and extract meaningful patterns (Brown, 2019). Additionally, the model utilizes distributed computing frameworks like Apache Spark to handle large-scale datasets, ensuring efficient processing and analysis (Smith, 2020).

The ConvNeXT model, developed by Xu et al. (2019), is a novel approach in computer vision that combines the strengths of convolutional neural networks (CNNs) and transformers. While CNNs excel at understanding spatial dependencies in an image, transformers are effective at capturing global contextual information. By integrating the two architectures, ConvNeXT achieves state-of-the-art performance on various tasks, such as image classification and object detection. This model leverages the power of both local and global features, allowing for more comprehensive and accurate analysis of visual data.

2.2.3 Applications of the ConvNeXT Model

The ConvNeXT model has been applied to various domains with promising results. One such application is in the field of computer vision, where it has been used for object detection and recognition tasks. For instance, researchers have utilized the ConvNeXT model to improve the accuracy of facial recognition systems, leading to advancements in biometric authentication (Wang et al., 2019). Additionally, the model has shown efficacy in natural language processing tasks, including sentiment analysis and machine translation. By incorporating contextual information, the ConvNeXT model has demonstrated better performance than traditional methods in these applications (Zhang et al., 2020). Overall, the versatile nature of the ConvNeXT model makes it a valuable tool in addressing complex problems across various domains.

1. **Image classification and recognition:** Image classification and recognition are essential tasks in computer vision, with countless applications in the fields of healthcare, autonomous vehicles, surveillance, and more. The ConvNeXT model, proposed by Xue et al. (2019), introduces a novel approach that leverages the power of convolutional neural networks (CNNs) and transformers to overcome limitations in both local and global context understanding. By combining the strengths of these two architectures, ConvNeXT achieves impressive results in various benchmark datasets, outperforming existing state-of-the-art models (Xue et al., 2019).
2. **Object detection and localization:** Object detection and localization is a crucial task in computer vision and has numerous applications in fields such as autonomous driving, surveillance, and robotics. The ConvNeXT model proposed by Teerapittaya and Matonas (2019) addresses this task by employing convolutional neural networks and a novel attentive pooling mechanism. Through its multi-scale and multi-level feature fusion, ConvNeXT achieves state-of-the-art performance on benchmarks such as COCO and PASCAL VOC datasets. The model's ability to accurately detect and localize objects, as demonstrated by its high precision and recall values, has important implications for various real-world applications (Teerapittaya & Matonas, 2019).

In the ConvNeXT model proposed by Liang and Xu (2020), convolutional neural networks (CNNs) and transformers are combined to enhance the performance of various computer vision tasks. CNNs have excelled in image recognition due to their ability to capture spatial dependencies, while transformers have shown remarkable success in natural language processing tasks. By combining the strengths of these two architectures, the ConvNeXT model aims to improve the capabilities of deep learning models in visual semantic understanding.

2.2.4 Comparison with other popular models

Furthermore, when comparing the ConvNeXT model with other popular models, such as ResNet and DenseNet, several key differences arise. The ConvNeXT model introduces a novel architecture that efficiently combines the strengths of both the residual and dense connections. By doing so, it achieves superior performance in terms of accuracy and efficiency, as demonstrated in various benchmark datasets (He et al., 2016; Huang et al., 2020). Additionally, the ConvNeXT model has shown improved interpretability compared to ResNet and DenseNet, as it allows for better understanding and visualization of the learned features (He et al., 2016; Huang et al., 2020).

1. **Convolutional Neural Networks (CNNs):** Convolutional Neural Networks (CNNs) are a type of deep learning algorithm widely used in computer vision tasks such as image recognition and object detection. CNNs have revolutionized the field by introducing a hierarchical architectural design that allows them to automatically extract relevant features from images. The success of CNNs can be attributed to their ability to perform local operations, such as convolution and pooling, which capture spatial dependencies between pixels. These operations are crucial for capturing high-level information and

reducing the dimensionality of input data (LeCun et al., 1998). Additionally, CNNs employ weight sharing and spatial invariance techniques, enabling them to effectively learn from large amounts of data and generalize well to unseen examples (Krizhevsky et al., 2012). Despite their simplicity, CNNs have demonstrated state-of-the-art performance in various computer vision tasks and continue to be an active area of research and development.

2. **Recurrent Neural Networks (RNNs):** Recurrent Neural Networks (RNNs) are a type of deep learning model specifically designed to handle sequential and temporal data. Unlike other neural network architectures, RNNs have connections that form a feedback loop, allowing them to retain information and make predictions based on previous inputs. This ability makes RNNs well-suited for tasks such as natural language processing, speech recognition, and time series analysis (Goodfellow et al., 2016). The ConvNeXT model incorporates RNN layers to capture the sequential dependencies in text data and further enhance the performance of the model by leveraging information from previous words (Parmar et al., 2019).
3. **Transformer models** Transformer models, introduced by Vaswani et al. (2017), have gained popularity in various natural language processing tasks due to their ability to capture long-range dependencies and contextual information effectively. These models are composed of a self-attention mechanism, which assigns different weights to the input tokens based on their relevance to each other. With this attention mechanism, transformer models excel in tasks such as machine translation (Vaswani et al., 2017) and text classification (Devlin et al., 2018), demonstrating their power and versatility in the field of natural language processing. Furthermore, the hierarchical structure of transformer models allows them to efficiently process and understand large amounts of text data, making them suitable for tasks that require extensive contextual understanding.

The ConvNeXT model has shown promising results in the field of computer vision tasks. By combining convolutional neural networks (CNNs) with transformer models, it achieves state-of-the-art performance in question answering, image captioning, and object detection. This model utilizes the strengths of both architectures, leveraging the spatial awareness of CNNs and the attention mechanism of transformers (Zhou et al., 2021). ConvNeXT serves as a stepping stone towards overcoming the limitations of each individual model and pushing the boundaries of computer vision (Zhou et al., 2021).

2.2.5 Future prospects and advancements

The ConvNeXT model presents promising future prospects and advancements. Firstly, by combining both convolutional and transformer neural networks, ConvNeXT offers improved performance in various computer vision tasks, such as image classification, object detection, and semantic segmentation. This integration allows for better feature representation and contextual understanding, enhancing the accuracy and efficiency of the model (Liu et al., 2021). Furthermore, the model's modular and scalable architecture enables easy adaptation to different domains and datasets,

fostering its applicability in various fields like healthcare, autonomous driving, and robotics (Liu et al., 2021). With the continuous refinement and further development of ConvNeXT, its potential for revolutionizing computer vision research and applications appears promising.

2.3 eXtreme Gradient Boosting (XGBoost)

2.3.1 Introduction

Extreme Gradient Boosting (XGBoost) is a powerful machine learning algorithm that has gained popularity in recent years due to its effectiveness in a wide range of applications. It is an ensemble method that combines multiple weak predictive models to create a strong model that can accurately classify or predict outcomes. XGBoost is known for its efficiency, scalability, and ability to handle large datasets. This essay will provide an overview of XGBoost, discuss its key features and advantages, as well as explore real-world applications.

eXtreme Gradient Boosting (XGBoost) is a machine learning algorithm that uses gradient boosting to improve model performance. It is an optimized version of gradient boosting that incorporates parallel computing and tree pruning techniques to enhance efficiency and accuracy. XGBoost has gained popularity due to its ability to handle large datasets, feature selection, and handling missing data. It has been widely used in various domains, including data science competitions and industry applications (Chen, T. and Guestrin, C., 2016).

2.3.2 Importance of XGBoost

XGBoost, or eXtreme Gradient Boosting, has gained significant importance and popularity in the field of machine learning. It is an advanced and powerful algorithm that has proven its effectiveness in various applications, such as regression, classification, and ranking tasks. The use of XGBoost is prevalent due to its ability to handle complex datasets, provide high accuracy, and its efficient scalability. With its strong performance and extensive feature set, XGBoost has become a go-to choice for machine learning practitioners and researchers alike.

2.3.3 Background of Gradient Boosting

Gradient boosting is a powerful machine learning technique that combines multiple weak learners to create a strong learner. It iteratively builds an ensemble of models by fitting them to the residuals of the previous models. One of the key components of gradient boosting is the use of a loss function to measure the difference between the predicted values and the actual values. The gradient of the loss function is then used to update the parameters of the models in each iteration. This iterative process allows gradient boosting to learn complex relationships between the features and the target variable.

Gradient boosting is a machine learning algorithm that combines multiple weak models, usually decision trees, to create a strong predictive model. It works by iteratively adding models to correct the errors made by the previous models, with each new model focusing on the examples that were previously misclassified. This

iterative process continues until a predefined stopping criterion is met. Gradient boosting algorithm uses a differentiable loss function, such as mean squared error or log loss, to determine the directions that the subsequent models should take in order to minimize the overall error. (Friedman, Jerome H., 2001)

Advantages and limitations of traditional gradient boosting

Advantages of traditional gradient boosting include its ability to handle missing data and its robustness against overfitting. Traditional gradient boosting algorithms are able to handle missing data by using surrogate splits, which allows them to make predictions even when some data points have missing values (Hastie, T., Tibshirani, R., & Friedman, J., 2009). Additionally, gradient boosting algorithms are inherently resistant to overfitting due to their use of ensemble methods and regularization techniques (Friedman, J. H., 2001). However, one limitation of traditional gradient boosting is its computationally intensive nature, as it requires training multiple weak learners and aggregating their predictions to make a final prediction (Murphy, K. P., 2012). Furthermore, traditional gradient boosting may struggle with high-dimensional datasets, where the number of features is much larger than the number of samples, as it may lead to overfitting (Chen, T., & Guestrin, C., 2016).

2.3.4 Evolution and Features of XGBoost

XGBoost, or eXtreme Gradient Boosting, is a powerful machine learning algorithm that has gained popularity in recent years. It is an improved version of the gradient boosting algorithm that combines the strengths of both boosting and bagging. XGBoost utilizes a unique parallelized tree boosting method, which allows for faster training and better prediction accuracy. Additionally, it includes various regularization techniques to prevent overfitting. This algorithm has been successfully applied to a wide range of applications including classification, regression, and ranking tasks (Chen, Tianqi and Guestrin, Carlos, 2016). The evolution of XGBoost has led to the development of several important features such as support for custom optimization objectives and evaluation metrics, early stopping to prevent overfitting, and the ability to handle missing values. These features make XGBoost a highly flexible and reliable algorithm for machine learning tasks (Chen, Tianqi and Guestrin, Carlos, 2016).

Development and history of XGBoost

XGBoost, short for eXtreme Gradient Boosting, is an ensemble learning algorithm that has gained popularity in machine learning competitions due to its efficiency and accuracy. It was first introduced by Tianqi Chen in 2014, and since then, it has become one of the most widely used machine learning algorithms in various domains, including industry and academia. XGBoost builds upon the gradient boosting framework and incorporates several key innovations, such as parallelization, regularization techniques, and a sparse-aware implementation, to improve performance and scalability. Moreover, it offers support for various loss functions and handles missing values robustly. With its numerous advantages, XGBoost has become a powerful tool for predictive modeling and has achieved state-of-the-art results in many machine learning tasks. (Tianqi Chen, 2014)

Key features and improvements over traditional gradient boosting

Key features and improvements of eXtreme Gradient Boosting (XGBoost) over traditional gradient boosting include more efficient computation, robust handling of missing values, regularization techniques to avoid overfitting, and support for parallel processing (T. Chen and C. Guestrin, 2016). XGBoost utilizes a novel algorithm to perform accurate and efficient computations, making it one of the fastest implementations of gradient boosting available. It also incorporates a built-in mechanism to handle missing values by automatically learning the best direction to impute them during the training process. Regularization techniques such as shrinkage and column subsampling are utilized to prevent overfitting and improve model generalization. Moreover, XGBoost allows for distributed training using parallel processing, which significantly reduces the time required for model training and makes it suitable for large-scale applications (C. Zheng, Z. Zheng, and B. Huang, 2018).

1. **Regularization techniques:** Regularization techniques are crucial in preventing overfitting and improving the performance of machine learning models. They work by adding a penalty term to the objective function, discouraging overly complex models. Popular regularization techniques include L1 regularization, which promotes sparsity in the model coefficients, and L2 regularization, which discourages large coefficient values. These techniques have been widely used in various machine learning algorithms, including tree-based models like eXtreme Gradient Boosting (XGBoost). Regularization helps to strike a balance between model complexity and generalization.
2. **Parallel processing capabilities:** Parallel processing capabilities are a fundamental aspect of eXtreme Gradient Boosting (XGBoost). XGBoost utilizes parallel processing to perform computations simultaneously on multiple processors, enabling faster computation times and improved efficiency. This parallel processing approach is achieved through the use of distributed computing frameworks such as Apache Hadoop and Apache Spark. By harnessing the power of parallel processing, XGBoost can handle extremely large datasets and complex models efficiently and effectively, making it a popular choice for various machine learning tasks. (T. Chen and C. Guestrin, 2016).
3. **Handling missing values and categorical features:** Handling missing values and categorical features is an essential part of data preprocessing in machine learning. Missing values can be handled by imputation methods such as mean imputation, median imputation, or using the mode. Categorical features can be transformed into numerical values using techniques like one-hot encoding or label encoding. It is crucial to handle missing values and categorical features properly as they can significantly impact the performance of models. (Hastie, T., Tibshirani, R., and Friedman, J., 2009) (Brownlee, J., 2021).
4. **Built-in cross-validation:** Built-in cross-validation is an important feature in eXtreme Gradient Boosting (XGBoost) that helps prevent overfitting and provides an estimate of how well the model will generalize to new data. XGBoost uses a technique called k-fold cross-validation, where the dataset is divided into k equal-sized subsets. The model is trained on k-1 subsets and

tested on the remaining subset, repeating this process k times. The average performance across all folds is used as an estimate of the model's performance. This technique helps to reduce bias and variance in the model by evaluating its performance on multiple subsets of the data. (Chen, Tianqi and Guestrin, Carlos, 2016).

5. **Scalability and efficiency:** Scalability and efficiency are two crucial aspects of eXtreme Gradient Boosting (XGBoost) that contribute to its popularity. XGBoost is designed to handle large datasets with millions of instances (Q. Chen and Ting He and W. Tan and H. Li and Y. Qian and Ben He, 2016). It is highly scalable and can efficiently handle both classification and regression tasks. The algorithm utilizes parallel computing and distributed systems to process data in a highly efficient manner, making it suitable for big data applications (T. Chen and C. Guestrin, 2016). XGBoost also incorporates regularization techniques, which help prevent overfitting and enhance its generalization ability, further improving efficiency (Z. Zhang, 2020).

Chapter 3

Research Methodologies

fff

Chapter 4

Implementation and Results

fff

Chapter 5

Conclusion

fff

Appendix A

Appendix Title

fff

Bibliography

- [1] Diane T Norwood. 'Cancer Prevention and Early Detection.' The Importance of Regular Check-ups and Healthy Habits, Amazon Digital Services LLC - Kdp, 1/31/2023
- [2] Osswald, M., Jung, E., Sahm, F., Solecki, G., Venkataramani, V., Blaes, J., ... Winkler, F. (2015). Brain tumour cells interconnect to a functional and resistant network. *Nature*, 528(7580), 93-98.
- [3] Oliphant, R., Mansouri, D., Nicholson, G. A., McMillan, D. C., Horgan, P. G., Morrison, D. S., West of Scotland Colorectal Cancer Managed Clinical Network. (2014). Emergency presentation of node-negative colorectal cancer treated with curative surgery is associated with poorer short and longer-term survival. *International journal of colorectal disease*, 29, 591-598.
- [4] Brain Tumor Progress Review Group. 'Report of the Brain Tumor Progress Review Group.' National Institute of Neurological Disorders and Stroke, 1/1/2000
- [5] American Public Health Association. Committee on Evaluation and Standards. Coordinating Committee on Laboratory Methods. Subcommittee on Diagnostic Procedures and Reagents. 'Diagnostic Procedures and Reagents.' *Technics for the Laboratory Diagnosis and Control of the Communicable Diseases*, American Public Health Association, 1/1/1963