

Research Journal of Pharmaceutical, Biological and Chemical Sciences

An Efficient Classifier Approach for Brain Tumor Prediction.

Sumathi A*, Poornima M, Kalaichelvi V, Manimozhi K and Meenakshi P.

SASTRA University, Kumbakonam, Tamil Nadu, India.

ABSTRACT

Decision tree induction algorithms have been used in many application areas one of these area is medical diagnosis. Several Decision tree algorithms have been used to classify the brain tumor and its types. A non-backtracking decision tree algorithm c4.5 adopts greedy approach in which decision trees are constructed in top down recursive divide and conquer manner. This algorithm predicts categorical class label with discrete attributes. But some of the cases it is hard to classify the subtypes because of distance based attribute and it leads to less prediction accuracy, which is the major drawback. To overcome this issue, this projected work proposed a new technique called Minimum Spanning Tree Classifier (MSTC) which improves the classification accuracy in terms of removing irrelevant features from distance valued attributes, which leads to achieve an efficient prediction. This proposed MSTC helps to predict the category of the brain tumor with accuracy and which leads to proper discretion for medical treatment.

Keywords: Decision Tree, Information Gain, C4.5, MSTC

*Corresponding author



INTRODUCTION

The term 'Brain tumor' is a disease which affects brain and central spinal canal. They are classified into primary and secondary, in which the primary brain tumor occurs due to environmental and genetic disorders and the secondary brain tumors [1], occurs due to cancer in other parts of the body spreading to the brain cells. Data Classification [2] is one of the familiar techniques which are used to categorize the large data base with the use of discrete attributes. The goal of classification is to accurately predict the target class for each case in the data. Decision trees (DT)[3] is a classification technique which provides possible solutions, based on sequence conditions of continues and discrete attributes. Decision tree [4] induction algorithms have been used in many application areas one of these area is medical diagnosis. Several Decision tree algorithms have been used to classify the brain tumor and its types. In the existing system c4.5 [5] demonstrates better accuracy for biological data classification. However this work has identified that the classification accuracy of c4.5 is not suitable for distance based attributes and some of the cases it is hard to classify the subtypes because of distance based attribute and it leads to less prediction accuracy.

METHODOLOGY

This work proposed an efficient classifier to classify brain tumor dataset. To achieve this, this proposed work is implemented c4.5[6] with minimum spanning tree approach called minimum spanning tree classifier which improved the classification accuracy in terms of removing irrelevant features from hard training dataset which leads to achieve an efficient prediction. It improves classification accuracy by find the false positive data at minimum level and it removed from the dataset and apply the greedy search for find data easily which helps to predict type of tumor in an efficient way. This proposed model was evaluated with tumor dataset. Patient age, gender, tumor location and tumor growth rate are taken as input.

Information Gain

This splitting criterion minimizes the information needed to classify the attributes and also minimizes the expected number of tests needed to classify the attributes. It calculates using the following formula

Gain(A)= Info(D)- Infox (D)

where Info(D) and Infox(D) are given as,

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2 p_i$$
$$Info_A(D) = \sum_{i=1}^{v} \left(\frac{|D_i|}{|D|}\right) \times Info(D_i)$$

The attribute which have high information gain is chosen as the splitting attribute and it is the root node. But Information Gain bias towards the attributes on large number of values.

C4.5 Classification

C4.5 algorithm was developed with the information gain attribute selection measure for construct decision tree for classification and the input data to be classified based on this method to find the type of tumor and calculated the classification accuracy using confusion matrix.

Proposed Methodology

MSTC Classification

The new technique called MSTC (Minimum Spanning Tree Classifier) algorithm was developed for classification. Minimum spanning tree is to start with an empty value and try to add minimum related values one at a time, always making sure that what is built remains acyclic. The resulting values are always is a subset of some minimum spanning tree. First find the some sure false positive values of dataset then remove that



data from the database and construct the decision tree based on c4.5 approach and apply minimum spanning tree approach for find particular growth rate from the dataset then predict the type of tumor and its states. Classification accuracy is calculated using confusion matrix **[Fig. 1]**.

MSTC Algorithm

Get training dataset D with n data object

if D is not empty If Dj is false data Remove Dj from D if attribute is discrete attribute attribute_list = attribute_list+attribute

Apply attribute selection method (D, attribute_list) Generate Decision tree(D, attribute_list)

repeat

Apply Minimum spanning tree in terminal node If (input_data== Dj) Predict entity for class variables Break;

Until j not greater than n



Fig.1 Proposed Methodology



Implementation



Fig.2 Upload dataset



Fig.3 Information Gain Calculation



C4.5 0	CLASSIFICATION	×			Contraction of the local division of the loc	and the second			
← → C 🗋 localhost5607/proj0814/C4algo.aspx 🟠 =									
🗋 lightnir	ngnewtab 🖪	4 Compan	ative Stud 🕒 Big Farm	🞯 Hot Ga	me 🗋 देवियों फेशन 웅 Measures of prompt 🤞	solved Bank Exam 💿 confusion matrix ex	🕵 Page N	ot Found - C	» 造 Other bookman
Root	: Gender	: Mes	el Barris man prosent harven harven harven harven harven harven harven harven	C	4.5 Algorithm	Classification	2 - 4 - 10 - 4 - 10 - 4 - 10	the or A many lyso A haroon many yet happen od acadent	
code	agegroup	gende	location	grate	type	ptype	state	Contraction of the second	
c72.5	Teenage	M	Astrocytes	0	Pilocytic Astrocytoma	Tumor of Neuroepithelial Tissue	Yes	ad Stoin	· She
c/1.0	Old age	M	Cerebral Cortax	0	Glioblastoma	Tumor of Neuroepithelial Tissue	Yes	there are the	1 D
C/5.1	Older	M	Cranial	0.21	Cranopharyngioma	Tumor of Seller region	No	1 through	a la
c/1.0	Child	M	Ventricle	0.65	Glioma maliganant, NOS	Tumor of Neuroepithelial Tissue	INO	more to put	
C93.9	Child	M	Choroid Plexus	0	Choroid Plexus Tumors	Tumor of Neuroepithelial Tissue	Yes	not to the A	
c/1.9	Teenage	M	Neoplasms	0.82	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	taaine -	fran
c/1.0	Middle age	M	Cerebral Cortax	0.14	Glioblastoma	Tumor of Neuroepithelial Tissue	Yes	3	
C30.0	Old age	M	nasal Cavity	0.12	Offactory Tumor	Tumor of Brain	INO	E DOM.	305
c/1.0	Middle age	M	Cerebral Cortax	0.3	Glioblastoma	Tumor of Neuroepithelial Tissue	Yes	there are h	Tip
06.0	Teenage	M	Mid Brain	0	Brain Stem Glioma	Germ Cell tumor	No	many types -	
c80.8	1 eenage	M	cauda equina	0.55	Neuronal and Mixed Neuronal Gilal Tumors	Tumor of Neuroepitnelial Tissue	NO	mine to aut	Ę.m
-71.0	Tanana age	M	Verebrai Cortax	0 1	Glioblastoma	Tumor of Neuroepineiial Tissue	I es	happon	
-75.1	Middle and	M	Disultante Claud	0.4	Turner Of the Bitsitery	Tumor of Neuroepinenal Tissue	NO V		to end
072.0	Old area	M	Final cord	0.23	Norra Shaeth Tumora	Tumor of Seller region	Ne	3	
075.3	Middle age	M	Dineal Gland	0.28	Tymor of The Dineal Region	Tumor of Neuroapithalial Tissue	Vec	E DAN	I.C.
071.9	Teenage	M	Neoplasms	0	Emproval types	Tumor of Neuroepithelial Tissue	Ver	redizion	in a start
072.5	A dult	M	Astroautos	0	Dilacitia Astrocitama	Tumor of Neuroepithelial Tissue	No	many lynes in	and the second sec
c72.5	Teenage	M	Astrocytes	0.92	Pilocytic Astrocytoma	Tumor of Neuroepithelial Tissue	Ves	the heroes	7
C05.0	Teenage	M	Lumphatic	0.35	Lizznhoma	Tumor of Lymphoma	Vec	happen	La contra de la co
c72.5	Child	M	Astrocutes	0.9	Pilocutic Astrocutoma	Tumor of Neuroepithelial Tissue	Ves	at St.A.	b en
012.3	Cime	141	risabcytes	0.5	r nocyue risu ocytoma	ranor or recubeptuicital rissue	105	taains -	braa

Fig.4 C4.5 Classification

C4.5	CLASSIFICATION	< ۱						-	Street State of State				- 0 ×
← →	C 🗋 loca	lhost:	5607/proj081	4/C4algo.a	spx								☆ 〓
🕒 lightni	ngnewtab a	A Com	parative Stud	🗋 Big Farm (🞯 Hot G	ame 🕒	देवियों फेशन 🔗 Measure	s of prompt 🥖	Solved Bank Exam	(1) confusion matrix ex	👷 Page	Not Found - C	» 📋 Other bookmarks
c/5.1	Adult	M	Hypophysi	IS	U	lumor	s of the Pituitary		1 umor of Selle	r region	Yes	S- non-	an -
c72.5	Adult	М	Astrocytes		0	Pilocy	ic Astrocytoma		Tumor of Neur	oepithelial Tissue	Yes	there are	in lite
c86.8	Adult	М	cauda equi	ina	0	Neuro	nal and Mixed Neurona	l Glial Tumors	Tumor of Neur	oepithelial Tissue	Yes	many lynes	100 m
c72.0	Adult	М	Spinal core	4	0	Nerve	Sheath Tumors		Tumors os cra	nial and Spainal Nerves	Yes	mine that and	of 1
c70.2	Adult	Μ	Cerebral N	/Ieninges	0	Menin	gioma		Tumor of meni	nges	Yes	happon	1
c94.1	Adult	М	Pons		0	Diffuse	Astroytoma		Tumor of Neu	oepithelial Tissue	Yes	ant with A.	to eat
C83.2	Middle age	М	Spinal Mer	ninges	0	Meser	chymal Tumors		Tumor of meni	nges	Yes	taaint -	wàai
c75.1	Middle age	М	Pituitary G	land	0	Tumor	s Of the Pituitary		Tumor of Selle	r region	Yes	and Contraction	and a second
c72.0	Middle age	М	Spinal core	đ	0	Nerve	Sheath Tumors		Tumors os cra	nial and Spainal Nerves	Yes	red Brain	. The
c71.0	Middle age	М	Cerebral C	Cortax	0	Gliobla	istoma		Tumor of Neur	oepithelial Tissue	Yes	there are the	Dr.
C83.2	Old age	М	Spinal Mer	ninges	0	Meser	chymal Tumors		Tumor of meni	nges	Yes	f thereast	of a
c75.1	Old age	М	Pituitary G	land	0	Tumor	s Of the Pituitary		Tumor of Selle	r region	Yes	mine at put	The second
c72.0	Old age	М	Spinal core	1	0	Nerve	Sheath Tumors		Tumors os cra	nial and Spainal Nerves	Yes	and State	bea
Num	ber of R	ecoi	idaa2672		hall		Watating 25		Waadini)			braaain) A.S.	Waa
Child	Node1 :	Loc	pation : Ne	eoplasms			Bred Bran		. Bred Bre Bore and Brean			Fred Sign	Ste
code	agegr	oup	gender	locatio	m	rate	type		ptyp	e	state	mine to just	e m
c71.9	Teenage		М	Neoplasms		0.82	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	Yes	happon	
c71.9	Teenage		М	Neoplasms		0	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	Yes	cat and A	to eat
c71.9	Child		М	Neoplasms		0.7	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	No	2	Alexandra and a second
c71.9	Child		М	Neoplasms		0.78	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	Yes	Sed Bian	
c71.9	Child		М	Neoplasms		0	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	No	there are the	200
c71.9	Child		М	Neoplasms		0	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	No	of the horas	71
c71.9	Teenage		М	Neoplasms		0	Empryonal tumors	Tumor	of Neuroepithelial Ti	ssue	No	mine to just	1 m
c71.9	Child		М	Neoplasms		0	Empryonal tumors	Tumor	of Neuroepithelial T	ssue	Yes	ent St. A.	beat
-71.0	Terrer		M	Magulanus		0	Engineeral burgers	т	CNI		Ver	aaans	braai *

Fig.5 C4.5 Classification



ISSN: 0975-8585

C4.5 C	LASSIFICATION	×				POST IN TRANSPORT			
< > (C 🗋 localhost	t:5607/proj08	814/C4algo.aspx						☆ 〓
🗋 lightning	gnewtab 🖪 A Con	nparative Stud	🕒 Big Farm 🛭 🞯 H	ot Game	े देवियों केशन 🔗 Measures d	of prompt 👝 Solved Bank Exam 🝈 confusion matrix ex	. 🧕 👷 Page	Not Found - C	» 📋 Other bookmarks
c71.9	Teenage	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	3	
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	Sect Prain	5
c71.9	Teenage	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	there are 1	20
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	of the horoset	J.
c71.9	Teenage	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	mine to pus	E.
c71.9	Teenage	М	Neoplasms	0.82	Empryonal tumors	Tumor of Neuroepithelial Tissue		and Street.	for each
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	dadini 📯	hrada
c71.9	Teenage	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	R DAN	ATT.
Child	Node2 : Ag	egroup :	Child	rate	of theread more of ust happen to eat	ne fare real happin	state	of the horizon mine of sur- happen to all of the	
c71.9	Child	М	Neoplasms	0.7	Empryonal tumors	Tumor of Neuroepithelial Tissue	No	L'	
c71.9	Child	М	Neoplasms	0.78	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	Fred Brain	. Su
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	No	many lyno in	- 2.
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	No	of the horses	74
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	happon	K
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	and she to	to east
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	aaams	Waa
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	the non	ar.
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	bure are	in the
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	of the horoes	ti ali
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	more the put	the second
c71.9	Child	M	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	happurs	

Fig.6 C4.5 Classification

lightnin	ignewtab	Comparative Stud	C Big Farm	Hot Game	ि देवियों फेशन 😕 Measures	of prompt	Page Not Found - C	» 🦳 Other bookmar
							in man and and	
c71.9	Child	M	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	No	- 10 m - 40
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes .	b ca
c71.9	Child	М	Neoplasms	0.6	Empryonal tumors	Tumor of Neuroepithelial Tissue	No	WOOD
c71.9	Child	Μ	Neoplasms	0.7	Empryonal tumors	Tumor of Neuroepithelial Tissue	No Done	in the second se
c71.9	Child	M	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes ned Dran	. Dre
c71.9	Child	M	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes more brief	4
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	No mene at put	2. m
c71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	No happen	
c71.9	Child	м	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	No idaain	haa
c71.9	Child	м	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	
71.9	Child	М	Neoplasms	0.8	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes Fred Room	S.
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes have are	- 70
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes of the hurses	910
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes mine and aut	R.
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes and Jr. A.	is en
c71.9	Child	М	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	bràai
c71.9	Child	M	Neoplasms	0	Empryonal tumors	Tumor of Neuroepithelial Tissue	Yes	
Numl Rule	er Of Re	cord 93 in of the base Gender 1	M and Lo	ocatic	bay and bay and bay and bay and bad bad bad bad bad bad bad bad bad ba	s and Age Group Child	Starting 19 Starting 19 Start	

Fig.7 C4.5 Classification





🗋 FINAL REPORT 🛛 🗙	opporter thanking the	
← → C 🗋 localhost:5607/proj08	814/freport.aspx	☆ =
🗋 lightningnewtab 🛛 a Comparative Stud	🕒 Big Farm 🞯 Hot Game 🕒 देवियों केशन 😒 Measures of prompt 🦱 Solved Bank Exam 🛞 confusion matrix ex 🤱 Page Not Found - C	» 📋 Other bookmarks
The Second Secon	CLASSIFICATION REPORT	That Designed
to ad Strange	n bagan bagan bada ya bada	b ad 4 braaan
C4.5 Algorith	CONFUSION MATRIX	
Status ; bor or	Tumor Start A	Died The
Growth : hove	0.8 d have d have	74
Tumor Type:	n Empryonal tumors haven	6. S.
Primary Type	: Tumor of Neuroepithenal issue	ts eat - brànain
Accuracy	68.0722891566265	
Error Rate Barrow	0.319277108433735	. Great
of the horses	Back a family family and the second s	and the second se
the phage	n Kanad Kana	and the second se
b and the second	l bad	te est -
Sted Sto	n Sted San Sted San Sted San	Bred
of thereas	a many log a start of the former log a start	af man
mine of put	n Kanan Kalin Kanan Kanan Kanan	mine
b at the	f bad and bad	b est = b datain

Fig.8 Classification Report of c4.5

MS	STC CLA	SSIFICATION	×				tion through the lot		States and States	- 0 - ×
← →	C	🗋 localho	ost:5607	/proj0814/mstc.as	рх					☆ 〓
🗋 light	tningnev	vtab 🖪 A C	omparati	ve Stud 🕒 Big Farm	🞯 Hot	t Game 📋 देवियों केशन 🔗 Measures of pron	npt 👝 Solved Bank Exam 🐧	Confusion matrix ex	. 🧕 Page Not Found - C	» 📋 Other bookmarks
	Root :	M	a a a		Cla	assification Al	lgorithm, P karan karan karan		Annon Ann- Annon A	A A A A A A A A A A A A A A A A A A A
		1.4	The	1.Brain	1	. Bred Brain	. Bred Brain		Sted Brain	. Du
	code	Child	M	cauda equina	grate 0	Neuronal and Mixed Neuronal Glial Tumors	Tumor of Neuroepithelial Tissu	ie No	of harang	
	C83.2	Middle age	М	Spinal Meninges	0	Mesenchymal Tumors	Tumor of meninges	No	to cal Ser A	ben
	C71.7	Teenage	M	Mid Brain	0.24	Brain Stem Glioma	Germ Cell tumor	No	Tanta Containe	
	C71.6	Middle age	М	Cerebellum	0.1	Medulloblastoma	Tumor of Neuroepithelial Tissu	ie Yes	· Bred String	. She
17 m	c72.5	Teenage	М	Astrocytes	0.92	Pilocytic Astrocytoma	Tumor of Neuroepithelial Tissu	ie Yes	many lines	
	c75.1	Middle age	М	Pituitary Gland	0.23	Tumors Of the Pituitary	Tumor of Seller region	Yes	nure part	E E
	C71.6	Middle age	М	Cerebellum	0.33	Medulloblastoma	Tumor of Neuroepithelial Tissu	ie No	bod St. A.	b ezi Vaz
	c70.2	Adult	М	Cerebral Meninges	0.89	Meningioma	Tumor of meninges	Yes	Statistic Contraction	
	c86.8	Adult	м	cauda equina	0.35	Neuronal and Mixed Neuronal Glial Tumors	Tumor of Neuroepithelial Tissu	ie Yes	Bred Stor	· Su
	c71.0	Old age	М	Cerebral Cortax	0.16	Glioblastoma	Tumor of Neuroepithelial Tissu	ie Yes	of the harden	- di
	c72.5	Adult	М	Astrocytes	0.95	Pilocytic Astrocytoma	Tumor of Neuroepithelial Tissu	ie Yes	mine to put	1
	C83.2	Old age	М	Spinal Meninges	0.6	Mesenchymal Tumors	Tumor of meninges	Yes	beat Str. A.	beat
	C30.0	Old age	М	nasal Cavity	0.21	Olfactory Tumor	Tumor of Brain	Yes	braaans -	braai -

Fig.9 MSTC Classification

ISSN: 0975-8585





Fig.10 Classificatiion Report of MSTC



Fig.11 Comparison of Classification Accuracy



Fig.12 Classification Accuracy of c4.5

2016

RJPBCS 7(4)



ISSN: 0975-8585



Fig.13 Classification Accuracy of MSTC

CONCLUSION

C4.5 is one of the familiar decision tree techniques. C4.5 algorithm performs well in constructing decision trees and extracting rules from the biological dataset. However, for biological dataset Minimum Spanning Tree Classifier algorithm is needed in order to provide better classification accuracy. Minimum spanning tree concept is used for finding data in dataset as much as faster than exiting system. From the experiment results, MSTC achieved high classification accuracy than c4.5 and it also gives less error rate. MSTC helps to predict the category of the tumor with accuracy and which leads to proper discretion for medical treatment.

REFERENCES

- [1] Eman M. Ali et al. International Journal of Computer Applications 2015; 131(2).
- [2] P Rajendran , M Madheswaran. Journal of Medical Systems 2012; 36(2):747-764.
- [3] Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers, 2008; 3rd reprint.
- [4] R Jin, and G. Agrawal, In Proc. of ACM SIGKDD, 2003; 571-576.
- [5] A Sumathi et al. Journal of Theoretical and Applied Information Technology 2012; 35(2):208-213.
- [6] P Shanthakumar and P Ganeshkumar. Computers & Electrical Engineering 2015; 45: 302-311