Advance in Electronic and Electric Engineering.
ISSN 2231-1297, Volume 3, Number 3 (2013), pp. 333-338

http://www.ripublication.com/aeee.htm

ANN-DT: An Algorithm for Extraction of Decision Trees from Artificial Neural Networks

Madhuri Jha

Student Mtech, Rcew, Jaipur.

Abstract

Althoughartificialneuralnetworkscan representavarietyofcomplexsystemswithahigh degreeofaccuracy,theseconnectionistmodelsare difficulttointerpret.Thissignif- icantlylimitsthe applicabilityofneuralnetworksinpractice, especially whereapremiumisplacedonthecomprehensibility or reliabilityofsystems.Anovelartificialneural-network de-cision tree algorithm(ANN-DT)is therefore proposed, which extracts binarydecisiontreesfroma trainedneuralnetwork.The ANN-DTalgorithmusesthe neural networkto generate outputsforsamples interpolatedfromthetrainingdataset.In contrastto existing techniques,ANN-DT canextractrulesfrom feedforwardneuralnetworks withcontinuous outputs. These rules are extractedfrom the neural network withoutmakingassumptionsabouttheinternalstructureoftheneuralnetwor korthefeaturesofthedata.Anovelattributeselectioncriterionbasedonasigni ficanceanalysisofthevariablesontheneural-networkoutput isexamined. It is shown to have significant benefits in certaincaseswhencomparedwiththestandardcriteria ofminimum weightedvarianceoverthebranches. In threecase studiesth eANN-DT algorithmcomparedfavorably with CART, astandarddecisiontreealgorithm.

Index Terms: Decision trees, hybridsystems, induction, neuralnetworks, ruleextraction, sensitivity analysis.

1. Introduction

Duringthelastdecadeinterestinartificialneuralnet- workshasgrownsignificantly, owingtotheirabilityto represent nonlinearrelationshipsthataredifficulttomodel bymeanso fothercomputationalmethods. Moreover,neural networksare easytoimplement,are robustunder the influence ofnoise,donotrequire*apriori*knowledgewithregardto thedistributionsofdata, andcanbeparallelizedwhererapid

computationiscritical. However, butforthesimpleststructures, neural-network models are notoriously difficultto interpret. For example, thefactthatneuralnetworkshavelargedegreesoffreedom intheassignmentofweights,canleadtoasituationwhere twocompletely differentsetsofweightscanyieldnearly identical outputs. This drastically complicates the analysis andcomparison ofsimilarprocessesthataremodeledor controlledbydifferentneuralnetworks. Theopacityofneural networkscanbeseenasamajorbarriertotheirimplementation in a number of fields, such as medicine and engineering wheremissioncriticalapplications demandahighdegreeof confidenceinthebehaviorofrelevantmodels.

2. Extraction of rules Fromneural Networks

Themethods thatareusedtoextractrulesfromneural networkscanbecategorizedas*decompositional,pedagogical,* and*eclectic*[5],[6],basedontheapproach usedtocharac- terizetheinternalmodelofthenetwork. Withpedagogical techniquestheneuralnetworkistreatedasa“blackbox,” i.e.,rulesthatmapinputstooutputsareextracteddirectly, evenformultilayered neuralnetworks.Forexample,Thrun’s method [11], [12] of validity-interval analysis (VIA) uses linearprogramming todetermineifasetofconstraintsplaced onanetwork’sactivationvaluesisconsistent. Insteadof approximatingtheactivation levelsofthehiddenunitsas thresholdfunctions, thelevelsareassumedtobeindependent ofoneanother.Sincethisassumption isnotalwaysappro- priate,thealgorithmdoesnotalwaysfind maximallygeneral rules. Other methods such as those developed by Craven andShavlik[6]andPop*et al.*[13]arecurrentlylimitedto discreteoutputs.Furthermore,CravenandShavlik’stechnique isimplemented usingeithertheVIanalysis[12]ortheKT methoddescribed byFu[8],wherebyitisimplicitly assumed thattheactivationsofthehiddenlayersareindependent,or thattheycanbetreatedasthresholdfunctions.

3. ANN-DTALGORITHM

ANN-DT(artificialneural-networkdecisiontree)algo- rithmdescribedinthispapercanbeseenaspedagogical, al- thoughitisnotboundbythelimitationsoftheabovestrategies.

Twovariantsofthealgorithmwereinvestigated, whichdiffer onlywithregardtothewayinwhichattributes andassociated splitpointsare selected.The proposedtechniquesare basedon thesamplingof aneural-networkrepresentationof the original data,insteadofattempting toextractrulesfromthesedataor interpretingtheinternalstructure oftheneuralnetwork. The ideaofsamplingtheneuralnetworkwasalreadyintroducedby CravenandShavlik[6]fordiscreteinputsandoutputsandwas extendedbyCravenandShavlik’sTREPAN[19]algorithmfor problemswithcontinuousand/ordiscreteinputs.TheTREPAN algorithm,which islimited todiscrete outputs,growsnodes in abestfirst order[19]andusesthegreedygainratiocriterion [3]toevaluateM-of-N splits.UnliketheTREPANalgorithm, ANN-DTcanbeappliedtodatasetswherebooththeinputsand outputscanassumediscreteorcontinuous values.Sincean earlierversionoftheANN-DTalgorithm restrictedtodiscrete output data has been published previously [20], [21], this paperfocusesspecifically onsystemswithcontinuousoutput variables.Moreover itisshownthatforproblems inwhich look-aheadcriteriaarerequired thenovelsignificanceanalysis proposedinthispapercanhaveappreciable advantagesover greedyattributeselectioncriteria.

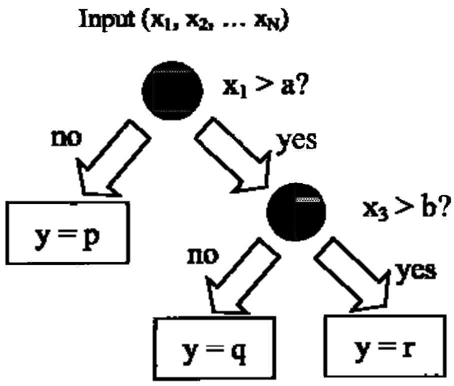

Fig. 1: Asimpleunivariatebinarydecisiontree.

AsisshowninFig.1,theANN-DTalgorithm generatesa univariatedecisiontreefromatrainedneuralnetwork.Agiven decisionnodeinthetreereturns “true”ifthenode’ssingle specified variableexceedsacertainthresholdand“false” otherwise.Foreachnodethealgorithm decidesonwhich variabletopartitionthesetofdata,afterwhichthethreshold ofthatvariablehastobedetermined.Morespecifically, the univariatedecisiontreeisconstructed byuseofthealgorithm toexaminetheresponses oftheneuralnetworkinthefeature spaceandtoconductasensitivityorsignificance analysisof thedifferentattributesorexplanatory variablespertainingto theseresponses.

4. TrainingoftheArtificial NeuralNetworks

TheANN-DTalgorithmisdepictedschematicallyinFig.2. Inthefirst stage of the procedure, aneural-network model capable of generalizingunderlyingtrendsinthedatahasto beconstructed. Inthisparticularinvestigation bothmultilayer perceptron[22] andradialbasisfunctionneuralnetworkswere used.Theformer wastrainedbyagradient descentmethod [23], inwhichtheerrorwaspropagated back ward through the network[24]. Training the radial basis function net work [25], [26]entailsfindingasuitablesetofbasisnodesandweights totheoutputlayer.Sincetheconnectingweightsoftheoutput layeroftheradialbasisfunction neuralnetwork canbefound bysolutionofaleastsquaresproblem,themostdifficulttask isfindingagoodsetofbasisfunctions.Inthisinvestigation anevolutionaryalgorithm[28]wasusedtofindtheoptimal setofbasisnodes,insteadofmorefrequently usedmethods suchas*k*-meansclusteringalgorithms[27].

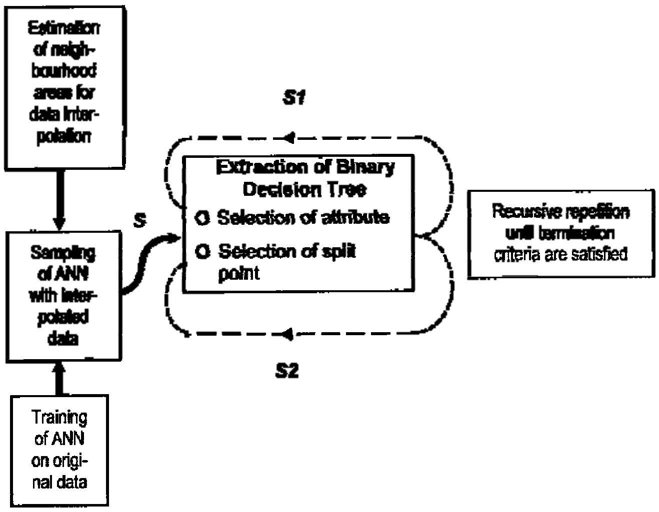

Fig. 2: Adiagrammaticre presentation of the algorithm to extractrulesfromanartificialneuralnetwork.

5. Conclusion

Anovelapproachhasbeendeveloped toextractdecision treesfromtrainedfeedforward neuralnetworks,regardlessof thestructures ofthesenetworks. Itwasfoundthatinsome casestheserulesweresignificantlymorerepresentativeofthe behavioroftheneuralnetworkthanrulesextracted fromthe trainingdataonly.

References

[1] R.Davis,B.G.Buchanan,andE.Shortliffe, “Productionrulesasarepresentationfora knowledge-based consultation program,”*Artificial Intell.*,vol.8,no.1,pp.15–45,1977.
[2] J.R.Quinlan,“Inductionofdecisiontrees,”*MachineLearning,*vol.1, pp. 81–106,1986.
[3] *C4.5:ProgramsforMachineLearning.* SanMateo,CA:Mor-gan Kaufmann,1993.
[4] L.Breiman,J.H.Friedman,R.A.Olshen,andC.J.Stone,*Classification andRegressionTrees.* NewYork:ChapmanandHall,1984.
[5] R.Andrews,J.Diederich,andA.B.Tickle,“Surveyandcritique of techniques for extractingrules from trainedartificialneuralnetworks,”*Knowledge-BasedSyst.,*vol.8,no.6,pp.373–383,1995.
[6] M.W.CravenandJ.W.Shavlik,“Usingsamplingand queriestoextract rules fromtrainedneuralnetworks,”in*Proc.11thInt.Conf.Machine Learning,* SanFrancisco, CA, 1994.
[7] G.TowellandJ.W.Shavlik,“Extractingrefined rulesfromknowledge
[8] Based neuralnet works,” *Machine Learning,*vol.13,pp.71–101,1993.[8] L.M.Fu,“Rulelearningbysearching onadaptednets,”in*Proc.9th Nat.Conf.ArtificialIntell.,*Anaheim,CA,1991,pp.590–595.
[9] S. I. Gallant,*NeuralNetworkLearningandExpertSystems.* Cam-bridge, MA:MITPress,1993.
[10] I.K.Sethi,“Neuralimplementationoftreeclassifiers,”*IEEETrans. Syst., Man, Cybern.,* vol.25, pp.1243–1249,1995.