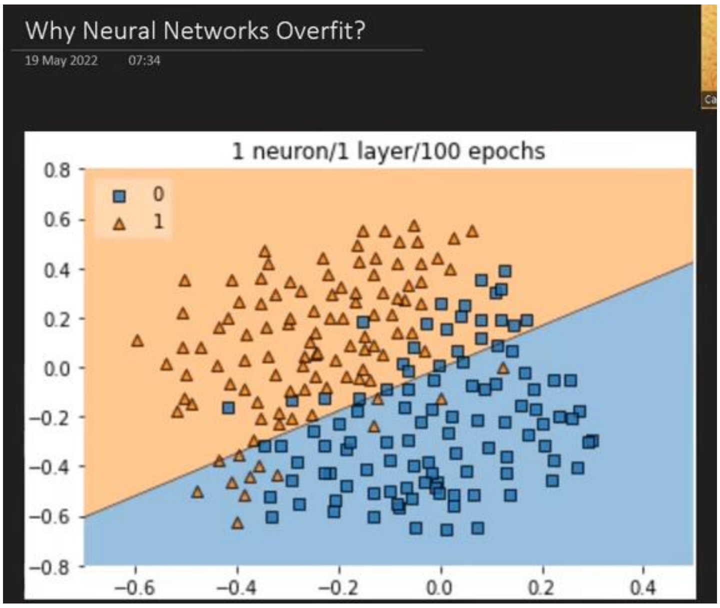
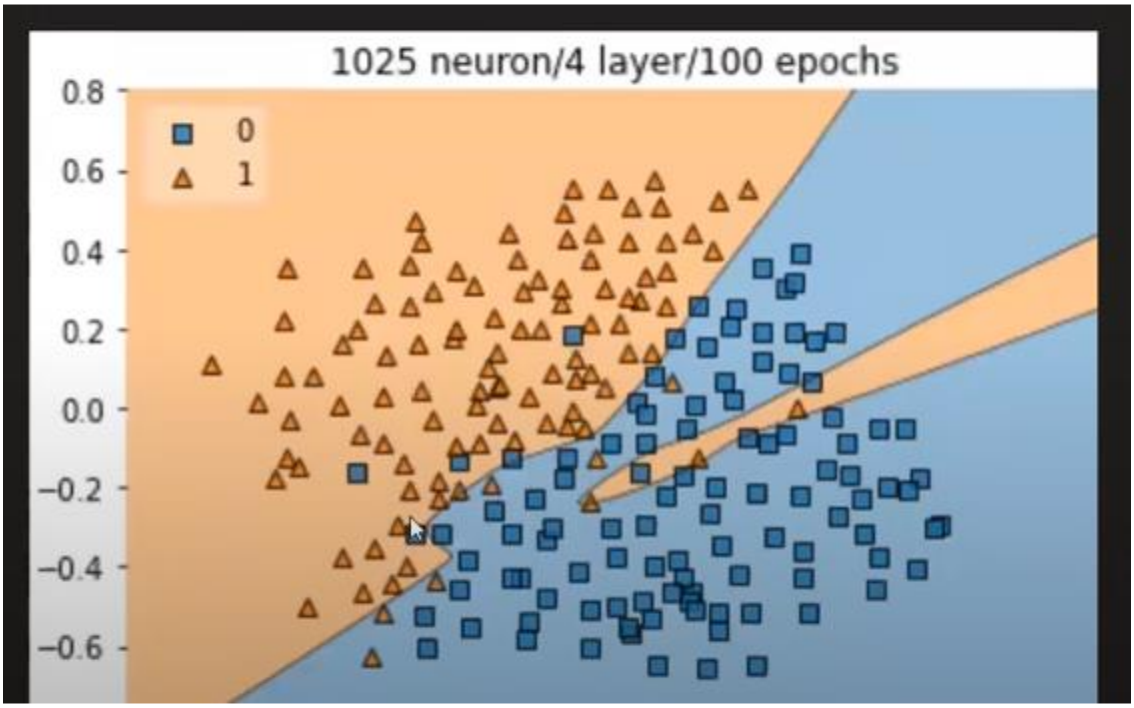
Regularization

NormalRevisiononOverfitting



ComplexityofModel



AsyoucanseeOverfithorha

Jitnazyaadaneuronsrahega,utnezyaadalinesrhengeutnazyaadavocapablehailinedrawkarkeclassifykrskta,&unnecessarypatternlele&ratherthanunderstandingtheessenceofData,voDatakorattlega.

NormalView

WaystoSolveoverfitting

1.) AddingMoreData
2.) ReducetheComplexityofModel

AddingMoreData

HerewecanaddmorerowsorwecandoDataAugmentation

Wecanproduceartificialsynteticdatafromgivendata

E.gDoghaithenuskolykleftketarafreversekrdenge,updownkrdengeesenayadatalaadenge

ReducingtheComplexityofModel

1.Dropout->BarbarharepochsmaiNeuronskogayabkarteraho

2.EarlyStopping->Itdetectkikahaseoverfittingstarthorha&stopatthatepochonly.

3.Regularization(Whichwewillstudy)

3TypesofRegularizationarefamous

L1,L2,&L1aurL2kaCombo

MostlyL2sehie99%kaamhojatahai,L1isnotusedmuch.

WhatisRegularization?

WehavetofindoutValueofWeights&Biasness.WefindoutbyminimizingLossFunction.SamewithLinearModel.

Regressionmai->MSE

ClassificationcasemaiBinaryCrossEntropy

→ min Loss function

$$C = \frac{1}{n}\sum_{i=1}^{n} L(y_i, \hat{y}_i)$$

Saareparameterkacostfunctionminimizekrtehai

SoInRegularization**Whatwedois,WeaddaPenaltyTerm.**

E.g

AsL2iscommoninDL,wewilltalkaboutit

$$C = L + \boxed{\frac{\lambda}{2n}\sum_{i=1}^{k} \|w_i\|^2} \rightarrow$$

LambdaisHyperparameter,jitnaweincreaseutnazyaadaoverfittingminimizehoga,&AgarBychancebhautzyaadaincreasekardengetouUnderfittingbhihosktahai.

InCaseofL1joSquarelagahaivonahihogabsssameformularhega

IntuitionBehindRegularization

Howweightsarereduced.

WhyRegularizationApplykrkegraduallyweights0keclosepemovekrnelagtehai&0kepasspochjat ahaibut0nahirehta.

L2Regularization->iskokbhikbhiweightdecaybhibulatehai

JitneepochschaltejaateutneWeightsdecreasehotejaate.

Summary

AsLossfunctionhighrehtatouvahioveffitrehtautna

Sooverfitreducekrnekeliyevogapkammkrtehum,lossfunctionkoreducekrte

&Assaareparametershummebadaweightderahehai

ThenweightkokammkrnekeliyePenaltyTermdaaltehai

RegularizationAddkrtehaihumthen

Weightkammhonelagtahai

L1Regularizationketime→Bhautsaareweights0hojatehai,sosparsemodelmiltahai,bhautnodes eliminatehojatahai

L2Regularizationketime->0kepassjaatahaibutkbhibhi0nahihotahai.

```
tion="relu",kernel_regularizer=tensorflow.keras.regularizers.l2(0.03)))
kernel_regularizer=tensorflow.keras.regularizers.l2(0.03)))
)
```

Insimpleterms,regularizationinthecontextofneuralnetworksisatechniqueused topreventamodelfrombecomingtoocomplexorfittingthetrainingdatatooclosely.Thegoalistoencouragetheneuralnetworktogeneralizewelltonew,unseendata.

Imagineyou'reteachingastudenttosolvemathproblems,andyougivethemaset ofpracticequestions.

Regularizationisliketellingthestudentnottomemorizetheanswersbuttounderstandtheunderlyingconcepts.Ifthestudentmemorizeseveryanswer,theymaynot beabletosolvenewproblemstheyhaven'tseenbefore.Similarly,inneuralnetworks,regularizationhelpspreventthemodelfrommemorizingthetrainingdatatooprecisely.

Therearedifferenttypesofregularizationtechniques,buttheyallaimtobalancethemodel'sabilitytolearnfromthedatawithoutoverfitting.Overfittingoccurswhen amodelbecomestootailoredtothetrainingdataandperformspoorlyonnew,unseendata.

Commonregularizationtechniquesinclude:

1. **L1andL2regularization:**Theseaddapenaltytermtotheneuralnetwork'slossfunctionbasedonthemagnitudesoftheweights.Thisencouragesthemodeltousesmallerweightsandpreventsanysingleweightfrombecomingtoodominant.
2. **Dropout:**Thisinvolvesrandomly"droppingout"(ignoring)someneurons duringtraining.Ithelpspreventthenetworkfromrelyingtoomuchonanyspecificsetofneuronsandencouragesamorerobustrepresentationofthedata.
3. **Earlystopping:**Thisinvolvesmonitoringthemodel'sperformanceonavalidationsetduringtrainingandstoppingthetrainingprocesswhentheperformancestartstodegrade.Thishelpspreventthemodelfromfittingthetrainingdatatooclosely.

Thesetechniqueshelpregularizethelearningprocess,makingtheneuralnetwork moreadaptableandbetterathandlingnew,unseendata.

L1andL2regularizationaretwocommontechniquesusedtopreventoverfittinginmachinel earningmodels,includingneuralnetworks.Theyworkbyaddingpenaltytermstothemodel'slossfunction,basedonthemagnitudesofthemodel'sweights.

1. **L1Regularization:**
   - AlsoknownasLassoregularization.
   - Involvesaddingtheabsolutevaluesoftheweightstothelossfunction.
   - Theregularizationtermisproportionaltothesumoftheabsolutevaluesoftheweights.
   - Encouragessparsityinthemodel,meaningittendstopushsomeweightstoexactlyzero.
   - Helpsinfeatureselectionbydrivingirrelevantorlessimportantfeatures'weightstozero.

   Mathematically,theL1regularizationtermisrepresentedas:
   L1RegularizationTerm=∑=1|| $\text{L1RegularizationTerm}=\lambda\sum_{i=1}^{n}|w_i|$
   Where w_i istheweightoftheithfeature,and λ controlsthestrengthoftheregularization.
2. **L2Regularization:**
   - AlsoknownasRidgeregularizationorweightdecay.
   - Involvesaddingthesquaredvaluesoftheweightstothelossfunction.
   - Theregularizationtermisproportionaltothesumofthesquaredvaluesoftheweights.
   - Encouragesthemodeltodistributetheweightmoreevenlyamongallfeatures.
   - Doesnotdriveweightstoexactlyzero,butreducestheirmagnitudes.

   Mathematically, theL2regularizationtermisrepresentedas:
   L2RegularizationTerm=∑=1 2 $\text{L2RegularizationTerm}=\lambda\sum_{i=1}^{n}w_i^2$
   Where w_i istheweightoftheithfeature,and λ controlsthestrengthoftheregularization.

Inbothcases,theregularizationstrength(λ)isahyperparameterthatneedstobetuned.Thehigherthevalueofλ,thestrongertheregularizationeffect.Regularizationhelpspreventoverfittingbypenalizingoverlycomplexmodelsandpromotingsimplermodelsthatgeneralizewelltonewdata.Manymachinelearningframeworksandlibrariesprovidebuilt-infunctionsforincorporatingL1andL2regularizationintoneuralnetworktraining.