Bayesian and Neural Networks for Motion Picture Recommendation

! "#$%!&'(#$)**%

+%*(%,#-%..&/&#0%,%'*#12&*3*
4563*%'7#8'%9&**%'#:&'/3%#4;#4.6<'&=
><"#?@A#@BBC

Abstract

123*#(2&*3*#<DD.3&*#E<F23,&#.&<',3,/#(&F2,3G)&*#(%#<#5<(<*&(#%9#E%63&*#5&*F'3!&5#!"#F%..<!%'<(36&#H*%F3<.I#<,5#F%,(&,(#<(('3!)(&*#3,#%'5&'#(%#F'&<(&#<#E3J&5#'&F%EE&,5&'#*"*(&E#9%'#E%63&*;#+<"&*3<,#,&(K%'L*A#(K%#6&'*3%,*#%9#,&)'<.#,&(K%'L*A#5&F3*3%,#('&&*A#<,5#*3ED.&#')..&#F.<**393&'* <'&#F%ED<'&5;#M(#3*#5&(&'E3,&5#(2<(#+<"&*3<,#<,5#,&)'<.,&(K%'L*#%)(D&'9%'E#(2&#'&E<3,3,/#(&F2,3G)&*;##+%(2#(&F2,3G)&*#<F23&6&5#'&E<'L<!."#23/2#(%DN?B#D'&F3*3%,A#<,#3ED%'(<,(#E&('3F#%9#'&F%EE&,5<(3%,#G)<.3(";#4,#<((&ED(#(%#F%,('<*(#'&F%EE&,5<(3%,#G)<.3("#9%'#F%,(&,(N%,."#<,5#F%..<!%'<(36&N%,."#5<(<*&(*#<*#F%ED<'&5#K3(2#<#5<(<*&(#5&*F'3!&5#!"#!%(2#F%,(&,(#<,5#F%..<!%'<(36&#<(('3!)(&*#"3&.5*#3,F%,F.)*36&#'&*).(*;#M,*)993F3&,(#F%,(&,(#3,9%'E<(3%,#3,#(2&#F)''&,(#5<(<*&(*#E<"#!&#(2&#'&<*%,#9%'#(23*;

Introduction

Definition: *A **recommender system** is a system that takes data about a user's past history in a certain industry, such as products they have purchased, movies they have seen, or websites they have visited, and predicts what the user may prefer to purchase or see in the future.*

: %E&#'&F%EE&,5&'#*"'*(&E*#<'&#*collaborative*#*"'*(&E*A#3,#K23F2#%(2&'#)*&'*0#D<*(#23*(%'3&*#<'&#)*&5#3,#('"3,/#(%#.3,L#(2&#D<'(3F).<'#)*&'#(%#<#/'%)D#%9#)*&'*#K3(2#*3E3.<'#3,(&'&*(*#%'#D)'F2<*&*;##123*#/'%)D#K3..#(2&,#3,9.)&,F&#K2<(#(2&#'&F%EE&,5&'#*"'*(&E#K3..#%)(D)(#!<*&5#%,#K2<(#(2&#/'%)D#.3L&*#%'#53*.3L&*;##P(2&'#*"'*(&E*#<'&#*content-based*"'*(&E*A#3,#K23F2#5&(<3.*#%9#(2&#D'%5)F(A#E%63&A#K&!*3(&A#%'#%(2&'#3(&E#<'&#F%ED<'&5#</<3,*(#(2%*&#%9#*3E3.<'#3(&E*#(2<(#(2&#)*&'#2<*#!&&,#3,#F%,(<F(#K3(2;##12&*&#*3E3.<'#3(&E*#<'&#2&.D&5#(%#/<)/&#K2&(2&'#(23*#)*&'#K3..#.3L&#%'#53*.3L&#(2&#3(&E;##:%E&#'&F%EE&,5&'#*"'*(&E*#)*&#<#E3J()'&#%9#F%..<!%'<(36&#<,5#F%,(&,(N!<*&5#<DD'%<F2&*;

$&F%EE&,5&'#*"'*(&E*#2<6&#!&F%E&#<#D%D).<'#*)!Q&F(#*(<'(3,/#3,#(2&#.<(&#?RRB0*H$&*,3FL#S#T<'3<,A#?RRUI;##><,"#%,.3,&#'&(<3.&'*A#*)F2#<*#4E<=%,;F%EA#)*&#'&F%EE&,5&'#*"'*(&E*#3,#%'5&'#(%#'&F%EE&,5#,&K#D'%5)F(*#(%#F)*(%E&'*#3,#%'5&'#(%#('"#(%#E<J3E3=&#D'%93(*;##V%'#&J<ED.&A#*)DD%*&#<#F)*(%E&'#!)"*#<#*F3&,F&N93F(3%,#!%%L;##12&#,&J(#(3E&#(2&#F)*(%E&'#63*3(*#(2&#*3(&A#3(#E3/2(#'&F%EE&,5#!%%L*#!"#(2&#*<E&#<)(2%'A#%'#E<"#*)//&*(#%(2&'#*F3&,F&N93F(3%,#!%%L*#(2<(#%(2&'#F)*(%E&'*#2<6&#!%)/2(#3,#<553(3%,#(%#(2&#%,&#(2&#F)*(%E&'#D)'F2<*&5#(2&#.<*(#(3E&#2&#%'#*2&#63*3(&5;##+"#2<63,/#(23*#("D&#%9#*"'*(&E#3,#D.<F&A#'&(<3.&'*#<'&#<!.&#(%#<..%K#F)*(%E&'*#(%#E%'&#&<*3."#93,5#D'%5)F(*#(2<(#(2&"#E<"#!&#3,(&'&*(&5#3,#(2<,#(2&3'#'&(<3.#*(%'&#F%),(&'D<'(*;

12&'&#<'&#*%E&#'&F%EE&,5&'#*"'*(&E*#K23F2#<..%K#)*&'*#(%#.3*(&,#(%#5399&'&,(#("D&*#%9#E)*3F#<,5#*&(#D'&9&'&,F&*#!<*&5#%,#(2&#*%,/*#(2&"#.3*(&,#(%;##12&*&#("D&*#%9#'&F%EE&,5&'#*"'*(&E*A#*)F2#<*#W<2%%;F%E0*#X4YZ-0F<*(#$<53%A#<..%K#<#)*&'#(%#/&(#3,*(<,(#)D5<(&*#3,#'&F%EE&,5<(3%,*#5)&#(%#!&3,/#<!.&#(%#'<(&#*%,/*#<*#(2&"#<'&#D.<"3,/;##4*#*)F2A#)*&'*#<'&#<!.&#(%#.3*(&,#(%#E%'&#*%,/*#(2<(#(2&"#F%).5#D%**3!."#.3L&#%,#<#D&'*%,<.3=&5#'<53%#*(<(3%,;

><,"#K&!*3(&*#<.*%#)*&#'&F%EE&,5&'#*"'*(&E*#(%#D&'*%,<.3=&#(2&3'#3,(&'9<F&#K3(2#)*&'*#3,#%'5&'#(%#E<3,(<3,#63*3(*#(%#(2&3'#*3(&;##V%'#&J<ED.&A#<(#<#,&K*#*3(&A#39#<#D&'*%,#&,(&'*#(2&3'#=3D#F%5&A#(2&#*3(&#E<"#F%,(<3,#.%F<.#,&K*A#K&<(2&'A#<,5#*D%'(*#(2<(#(2&#)*&'#E<"#D'&9&'#(%#'&<5#%6&'#,<(3%,<.#%'#/.%!<.#F%,(&,(;##+"#<..%K3,/#(23*#F)*(%E3=<(3%,#)*&'*#E<"#!&((&'#&,Q%"#(2&#*3(&#<,5#E%'&#&<*3."#93,5#<'(3F.&*#(2<(#(2&"#<'&#3,(&'&*(&5#3,;

>"#3,(&,(3%,#3*#(%#93,5#<#!&((&'#<./%'3(2E#K23F2#F%E!3,&*#F%,(&,(N!<*&5#5<(<

Machine Learning Concepts

><F23,&#.&<',3,/#3*#<#'<D35."#/'%K3,/#93&.5#K3(23,#F%ED)(&'#*F3&,F&;

Definition: "*A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" (Mitchell, 1997).*

Definition: *A **dataset** is a group of data. It is basically an n-by-m matrix with n rows and m columns. The rows are called **instances**. Instances are basically different occurrences of a situation. The columns are called **attributes**. Attributes are certain details that were recorded during every instance.*

V3/)'&#?#3*#K2<(#<#/&,&'3F#5<(<*&(#E<"#.%%L#.3L&;

Figure 17#Sample Dataset

Student	Graduation Year	Major	GPA

Decision Trees

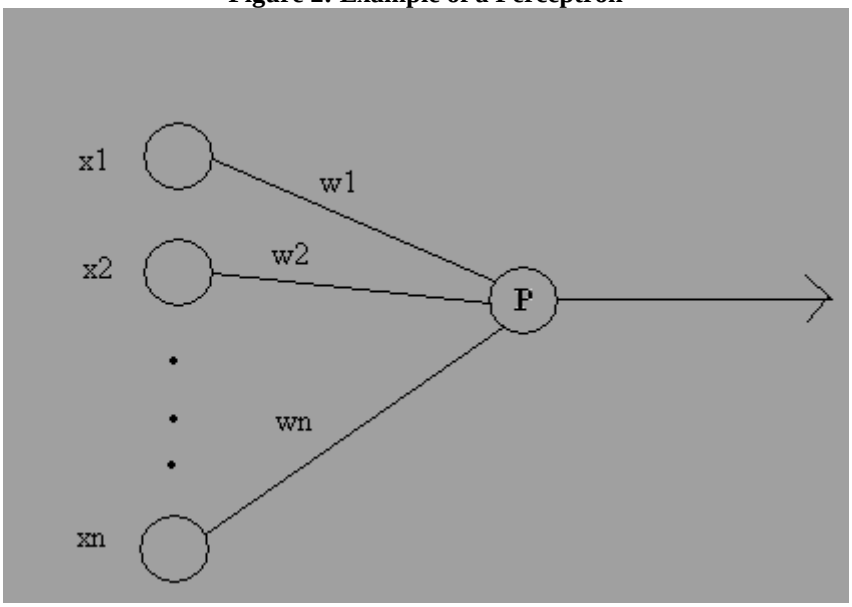
4 , %(2&'#F.<**393&'#K23F2#!)3.5*#)D% , #') .&#F.<**393&'*#3*#% , &#K3(2#5&F3*3% , #('&&*;## ` &F3*3% , #('&&*#!<*3F<..''#(<L&#<#.%%L#<(#E).(3D.&#<(('3!)(&*#< , 5#! '&<L#(2&E#3 , (%#*&D<'<(&# D<3'3 , /*A#D'&53F(3 , /#5399&'& , (#6<.)&*#9%'#(2&#(<' /&(#<(('3!)(&#! <*&5#% , #(2&#D<3'3 , /#%9#%(2&'# 6<'3<! .&*;##Z&K#! '< , F2&*#<'&#E<5&#9%'# , &K#D<3'3 , /*#< , 5#<*#<#'&*) .(#< , #& , (3'&#('&&#3*#E<5&A# K23F2#(2&#F.<**393&'#! '< , F23 , /#<(#&<F2#D%3 , (#5&D& , 53 , /#% , #K2<(#(2&#6<.)&#%9#<#D<'(3F) .<'# <(('3!)(&#3*;## > < , ''#<. /%'3(2E*#2<6&#! && , #E<5&#)*3 , /#5&F3*3% , #('&&*A#(2&#E%*(#! <*3F#<E% , /# (2&E#! &3 , /#(2&#M ` ^#<. /%'3(2E#H a)3 , .< , A#?RR^I;##0%K&6&'A#5&F3*3% , #('&&*#3 , 3(3<..''#2<5#<# D'%! .&E#3 , #%6&'93((3 , /#5<(<A#3 , #(2<(#'< , 5%E#%FF) ''& , F&*#K%) .5#F'&<(&# , &K#! '< , F2&*#(%# ('&&*#< , 5#K%) .5#(2&'&9%'&#.&<5#(%#<#E%'&#F%ED.&J#('&&A#% , &#(2<(#F%) .5#D'&53F(#9<.*&# %)(F%E&*#5)&#(%#*%E&#3 , *(< , F&*#K2%*&#(<' /&(#<(('3!)(&#6<.)&#K%) .5#! &#5399&'& , (#(2< , # %(2&'#3 , *(< , F&*#K3(2#*3E3.<'#6<.)&*#9%'#3(*#<(('3!)(&*;##4*#<#'&*) .(A# , &K#<. /%'3(2E*#K&'&# E<5&#)*3 , /#D') , 3 , /;##8') , 3 , /#K%) .5#'&E%6&#! '< , F2&*#%9#(2&#('&&#39#3(#3ED'%6&5#(2&# %6&'<..#<FF)'<F''#%9#(2&#('&&;##bJ<ED.&*#%9#5&F3*3% , #('&&#<. /%'3(2E*#3 , F.)5&# - \; [#H a)3 , .< , A# ?RR^I#K23F2#!)3.(#)D% , #M ` ^#! ''#<553 , /#D') , 3 , /A#<*#K&..#<*#<..%K3 , /# ,)E&'3F#<(('3!)(&# 6<.)&*;

Neural Networks

4 '(393F3<.# , &)'<.# , &(K%'L*#<'&#< , %(2&'#(''D&#%9#E<F23 , &#.&<' , 3 , /#(&F2 , 3G)&;##M(#K<*# 3 , *D3'&5#! ''#(2&# , &)'% , *#3 , *35&#<#2)E< , #! '<3 , #K23F2#F% , , &F(#(%#&<F2#%(2&'#< , 5#/& , &'<(& %)(D)(*#! <*&5#% , #*(3E).3#9'%E#%(2&'# , &)'% , *;##M , #<'(393F3<.# , &)'<.# , &(K%'L*A#D&'F&D('% , *# <'&#)*&5#3 , *(&<5#%9# , &)'% , *;##4#D&'F&D('% , #(<L&*#3 , #E< , ''#3 , D)(*#< , 5#<**3/ , *#*%E&# F% , *(< , (#F<..&5#(%#&<F2#% , &A#F<..&5#<#K&3/2(;##12&#K&3/2(#'&D'&*& , (*#(2&#3ED%'(< , F&#%9#(2&# 3 , D)(#(%#(2&#D&'F&D('% , ;## , #(2&#E<F23 , &#.&<' , 3 , /#F<*&#(2&#3 , D)(*#<'&#(2&#5399&'& , (#6<.)&*#%9# (2&#<(('3!)(&*#%9#< , #3 , *(< , F&;##12&#D&'F&D('% , #(2& , #<55*#<..#%9#(2&#3 , D)(*#E).(3D.3&5#! ''# (2&3'#K&3/2(*;##M(#(2& , #<DD.3&*#<#(2'&*2%.5#(%#(23*#6<.)&;##M9#3(#3*#<! %6&#BA#(2&#D&'F&D('% , #K3..# %)(D)(#<#?;##M9#(2&#6<.)&#3*# , %(#<! %6&#BA#(2& , #(2&#D&'F&D('% , #K3..#%)(D)(#<#N?;##V%'#&J<ED.&A# V3/)'&#@#3*#<#53</'<E#%9#<#D&'F&D('% , ;##

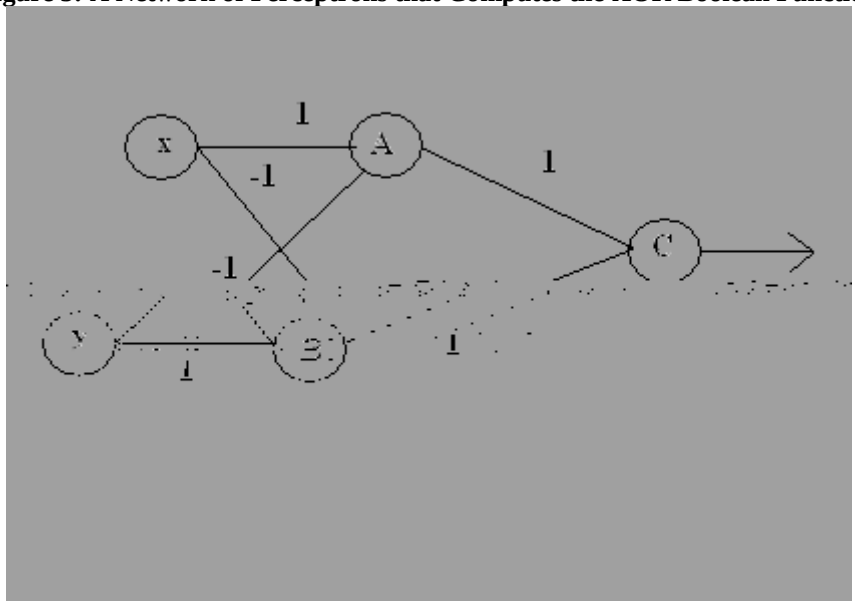
Formula: c36& , #3 , D)(*#J?A#J@A# d A#J¸A#<#D&'F&D('% , #K3..#<**3/ , #K&3/2(*#K?A#K@A# d A#K¸#(%# &<F2#%9#(2&E;##M(#K3..#(2& , #F<.F).<(&#J?K?#e#J@K@#e# d #e#J¸K¸#I#< , 5#(2& , #%)(D)(#<#*3 , /.&# 6<.)&#!<*&5#% , #K2&(2&'#%'# , %(#(2<(#6<.)&#3*#/'&<(&'#(2< , #B;

Figure 2: Example of a Perceptron



8&'F&D('%,*#F<,#!&#2%%L&5#)D#K3(2#%,&#<,%(2&'#.3L&#,&)'%,*A#3,#%'5&'#(%#F'&<(&#,&(K%'L*#F%,*3*(3,/#%9#E).(3D.&#.<"&'*#%9#D&'F&D('%,*;##V%'#&J<ED.&A#*)F2#<#,&(K%'L#F<,#!&#)*&5#(%#F%ED)(&#(2&#+%%.&<,#.%/3F#9),F(3%,#fP$;##V3/)'&#^#3*#<#53</'<E#%9#2%K#5%3,/#J#fP$#"#K%).5#K%'L;

Figure 3: A Network of Perceptrons that Computes the XOR Boolean Function



M,#(23*#F<*&A 4A#+A#<,5#-#<'&#D&'F&D('%,*;##4#K%).5#%,."#%)(D)(#<#?#39#J#3*#?#<,5#"#3*#B;##+#K%).5#%,."#%)(D)(#<#?#39#J#3*#B#<,5#"#3*#?;##-#K%).5#%,."#%)(D)(#<#?#39#4#%'#+#3*#%)(D)((3,/#<#?;##12&'&9%'&A#(23*#,&(K%'L#%9#D&'F&D('%,*#E%5&.*#(2&#fP$#9),F(3%,;##M,#/&,&'<.A#D&'F&D('%,*#<'&#'&9&''&5#(%#<*#,%5&*A#<,5#D&'F&D('%,*#(2<(#%,."#%)(D)(#6<.)&*#(2<(#/%#3,(%#E%'&#D&'F&D('%,*#<'&#F<..&5#2355&,#,%5&*;

Network Training

12&#K<"#<#,&)'<.#,&(K%'L#.&<',*#3*#!"#('<3,3,/#3(*#K&3/2(#6<.)&*;##b<F2#(3E&#<,#3(&'<(3%,#%9#D<**3,/#(2&#3,D)(*#(2'%)/2#(2&#,&(K%'L#3*#5%,&A#(2&#%)(D)(#6<.)&*#<'&#F%ED<'&5#</<3,*(#(2&#(<'/&(#<(('3!)(&0*#6<.)&*;##12&,#<#F<.F).<(3%,#3*#D&'9%'E&5#(%#F'&<(&#(2&#5&.(<#(2<(#E)*(#!&#<55&5#(%#(2&#%'3/3,<.#K&3/2(;##123*#F%,(3,)&*#),(3.#(2&#K&3/2(*#5%#,%(#F2<,/&A#%'#),(3.#(2&#<E%),(#%9#D'&N5&(&'E3,&5#3(&'<(3%,*#3*#'&<F2&5;##12&'&#<'&#*&6&'<.#K<"*#(%#)D5<(&#K&3/2(*#9%'#D&'F&D('%,*;##P,&#K<"#3*#(%#<**3/,#'<,5%E#K&3/2(*#(%#&<F2#3,D)(#<,5#(2&,#3(&'<(3,/#(2&#D&'F&D('%,#)*3,/#(2&#9%..%K3,/#').&#9%'#)D5<(3,/#&<F2#K&3/2(#K3#9%'#3,D)(#J37

Formula: K3 ←#K3#e#æw3#####K2&'&#####æw3 = ɖ(t g#%IJ3

M,#(2&#<!%6&#9%'E).<A#ɖ is the learning rate used to determine to what extent the K&3/2(#K3..#!&#F2<,/&5A#(#3*#(2&#(<'/&(#<(('3!)(&0*#6<.)&A#%#3*#(2&#%)(D)(#6<.)&#/36&,#!"#(2&#D&'F&D('%,A#<,5#J3#3*#(2&#3,D)(#6<.)&#9&5#3,(%#(2&#D&'F&D('%,;##123*#E&(2%5#/)<'<,(&&*#F%,6&'/&,F&#39#(2&#F.<**&*#<'&#.3,&<'."#*&D<'<!.&#H>3,*L"#S#8<D&'(A#?RC[I;##0%K&6&'A#E%*(#5<(<*&(*#5%#,%(#2<6&#.3,&<'."#*&D<'<!.&#F.<**&*;

4#*%.)(3%,#(%#(23*#D'%!.&E#K<*#)*3,/#&''%'#!<FLD'%D</<(3%,;##M,#(23*#K<"A#(2&#%)(D)(*#<'&#F<.F).<(&5A#<,5#(2&,#(2&#&''%'#3*#*&,(#!<FLK<'5*#(2'%)/2#(2&#,&(K%'L#(%#)D5<(&#(2&#K&3/2(*;##b<F2#%)(D)(#D&'F&D('%,#L#F<.F).<(&* its error term ŭL#)*3,/#(2&#9%..%K3,/#9%'E).<7

Formula: ŭL ← %LH?#g#%LIH(L g#%LI

0&'&A#%L#3*#(2&#%)(D)(#6<.)&#9%'#(2&#D&'F&D('%,A#<,5#(L#3*#(2&#(<'/&(#<(('3!)(&#6<.)&;##12&*&#6<.)&*#<'&#)*&5#3,#F<.F).<(3,/#(2&#&''%'#(&'E#9%'#&<F2#2355&,#,%5&#2#)*3,/#(2&#9%..%K3,/#9%'E).<7

Formula: ŭ2 ←#%LH?#g#%L)Ɇ wL2ŭL

Z%(&#(2<(#KL2#5&,%(&*#(2&#K&3/2(#9'%E#,%5&#2#(%#,%5&#L;#12&#(&'E#Ɇ wL2ŭL 9%'#2355&,#,%5&#2#3*#(2&'&9%'&#(2&#*)E#%9#(2&#K&3/2(*#(3E&*#(2&#&''%'#6<.)&*#%9#<..#(2&#%)(D)(#,%5&*#L#(2<(#<'&#F%,,&F(&5#9'%E 2;

12&#K&3/2(*#<'&#(2&,#)D5<(&5#3,#(2&#9%..%K3,/#K<"7

Formula: KQ3 ←#KQ3#e#æwQ3######K2&'&#######æwQ3 = ɖŭQJQ3

123*#)D5<(&#').&#'&5)F&*#(2&#E&<,#*G)<'&#&''%'#<(#(2&#%)(D)(#.<"&';##-%,,&F(3,/#E<,"#D&'F&D('%,*#(%/&(2&'#K3(2#E<,"#2355&,#.<"&'*#)*3,/#&''%'#!<FLD'%D</<(3%,#F<,#2&.D#3ED'%6&#<FF)'<F"#!)(#.&<5*#(%#<#*2<'D#3,F'&<*&#3,#('<3,3,/#(3E&;##T<'"3,/#(2&#(2'&*2%.5A#(2&#,)E!&'#%9#.<"&'*A#<,5#(2&#<E%),(#%9#(3E&#(<L&,#(%#('<3,#(2&#,&(K%'L#F<,#<..#2&.D#3ED'%6&#<FF)'<F";##1&*(*#2<6&#!&&,#5%,&#)*3,/#(2&*&#5399&'&,(#6<'3<(3%,*;H$)E&.2<'(A#h35'%KA#S#X&2'A#?RR\I;

Y*3, /#&''%'#!<FLD'%D</<(3%,#3,#(2&#K<"#<!%6&#)*&*#<#(2'&*2%.5#3,#%'5&'#(%#F%,6&'/&#(%#%,&#6<.)&;##0%K&6&'A#(2&'&#3*#<#K<"#(%#E%539"#&''%'#!<FLD'%D</<(3%,#K2&'&#3,*(&<5#%9#)*3,/#(2&#(2'&*2%.5A#<#D'%!<!3.3("#3*#3,*(&<5#F%ED)(&5A#<,5#(2<(#3*#)*&5#3,#5&(&'E3,3,/#(2&#%)(D)(#<(('3!)(&#6<.)&;##M,#(23*#K<"A#(2&#6<.)&#9%'#(2&#&''%'#(&'E#9%'#&<F2#%)(D)(#,%5&#L#K%).5#(2&,#!&#F<.F).<(&5#3,#(2&#9%..%K3,/#K<"7

Formula: ŭ$_L$ ←#%$_L$H($_L$ g#%$_L$I

12&#6<.)&#9%'#(2&#&''%'#(&'E#9%'#&<F2#2355&,#,%5&#2#K%).5#(2&,#!&#F<.F).<(&5#3,#(2&#9%..%K3,/#K<"7

Formula: ŭ$_2$ ←#%$_L$Ɇ w$_{L2}$ŭ$_L$

M,#(23*#K<"A#(2&#(2'&*2%.5#K%).5#!&#'&E%6&5#<,5#!&#'&D.<F&5#K3(2#<#D'%!<!3.3(";##T<'3%)*#K%'L#2<*#!&&,#5%,&#)*3,/#(23*#E%5393&5#9%'E#<*#K&..#H+3*2%DA#?RRCI

Bayesian Techniques

4,%(2&'#("D&#%9#F.<**393&'#)*&*#+<"&*3<,#'&<*%,3,/;##+<"&*3<,#(&F2,3G)&*#<'&#!<*&5#%,#D'%!<!3.3("#53*('3!)(3%,*#<,5#(2<(#)*3,/#(2&*&#D'%!<!3.3(3&*#%,#%!*&'6&5#5<(<#F<,#3ED'%6&#D&'9%'E<,F&;##M(#('3&*#(%#D'%5)F&#(2&#!&*(#2"D%(2&*3*#9'%E#*%E&#*D<F&#%9#2"D%(2&*&*#0#/36&,#*%E&#('<3,3,/#5<(<#`;#>%*(#%9#+<"&*3<,#.&<',3,/#'&.3&*#%,#+<"&*0(2&%'&E;#+<"&*0#(2&%'&E#<**)E&*#(2<(#9%'#&<F2#2"D%(2&*3*#2A#(2&'&#3*#<#D'3%'#D'%!<!3.3("#<.'&<5"#F<.F).<(&5A#F<..&5#8H2I;##8H`I#3*#(2&#D'3%'#D'%!<!3.3("#(2<(#(2&#('<3,3,/#5<(<#`#K3..#!&#%!*&'6&5;##8H`i2I#3*#(2&#D'%!<!3.3("#(2<(#(2&#('<3,3,/#5<(<#`#K3..#!&#%!*&'6&5#/36&,#(2<(#(2&#2"D%(2&*3*#2#2%.5*;##4,5#8H2i`I#3*#(2&#D'%!<!3.3("#(2<(#(2&#2"D%(2&*3*#2#K3..#2%.5#/36&,#(2&#('<3,3,/#5<(<#`;##1%#5&(&'E3,&#8H2i`IA#+<"&*0#(2&%'&E#D'%635&*#(2&#9%..%K3,/#9%'E).<7

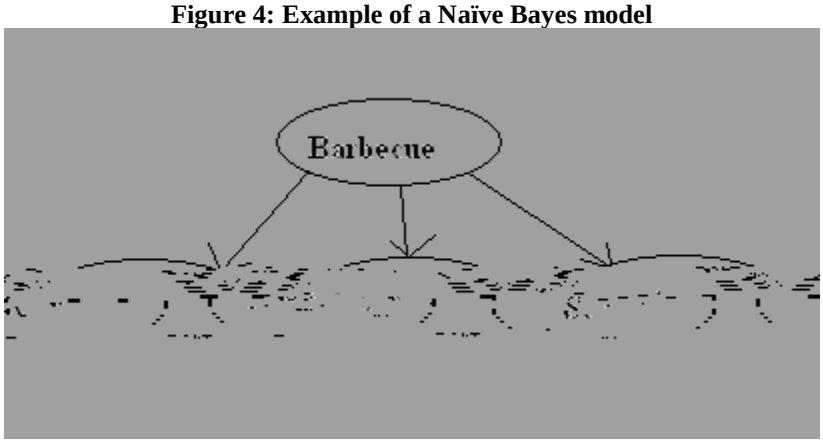
Formula: 8H2i`I#j#8H`i2I8H2I k 8H`I

+<"&*0#(2&%'&E#<..%K*#F.<**393F<(3%,#!"#*&.&F(3,/#(2&#E<J3E)E#<#D%*(&'3%'3#H>48I#2"D%(2&*3*;##V%'#&6&'"#2#3,#0A#8H`i2I8H2I#3*#F<.F).<(&5;##12&#5&,%E3,<(%'#8H`I#3,#+<"&*0#9%'E).<#3*#'&E%6&5#!&F<)*&#(2&#('<3,3,/#5<(<#`#,&6&'#F2<,/&*#9%'#&<F2#F2<,/&#%9#2#3,#0;##12&#2"D%(2&*3*#2#K3(2#(2&#23/2&*(#6<.)&#%9#8H`i2I8H2I#3*#(2&,#)*&5#(%#F.<**39"#&<F2#3,*(<,F&;

Naïve Bayes

12&'&#<'&#(K%#*(<,5<'5#E&(2%5*#3,#K23F2#+<"&*3<,#.&<',3,/#3*#5%,&;##P,&#3*#<#(&F2,3G)&#F<..&5#Z<I6&#+<"&*;##M,#Z<I6&#+<"&*A#(2&#<./%'3(2E#F'&<(&*#<#*&(#%9#<..#D%**3!.&#(<'/&(#<(('3!)(&*;##M(#(2&,#F<.F).<(&*#(2&#D'%!<!3.3("#(&'E*#8H2I#<,5#8H`i2I#<*#*(<(&5#<!%6&;##0%K&6&'A#3(#!'&<L*#&<F2#5399&'&,(#<(('3!)(&#6<.)&#5#%9#(2&#('<3,3,/#5<(<#`#<,5#F<.F).<(&*#&<F2#%,&#%9#(2&E#*&D<'<(&.";##M(#(2&,#(<L&*#(2&#E<J3E)E#6<.)&*#%9#8H2I#E).(3D.3&5#!"#(2&#D'%5)F(#%9#<..#(2&#D'%!<!3.3(3&*#%9#8H`i2I;#V%'#&J<ED.&A#F%,*35&'#V3/)'&#\;##

Figure 4: Example of a Naïve Bayes model



+<'!&F)&#3*#(2&#(<'/&(#<(('3!)(&A#<,5#2<*#(2&#6<.)&*#"&*#%'#,%;#12&#F.<**393&'#K%).5#93'*(#!)3.5#)D#(2&#5399&'&,(#3,*(<,F&*#%9#(2&#!<'!&F)&#5<(<#`#<,5#F<.F).<(&#&<F2#D'%!<!3.3("#/36&,#&3(2&'#"&*#%'#,%;#M(#K%).5#<.*%#F<.F).<(&#(2&#%6&'<..#D'%!<!3.3("#%9#!<'!&F)3,/;##V%'#&J<ED.&A#(%#F.<**39"#(2&#3,*(<,F&#H:),#j#%)(A#$<3,#j#,%,&A#1&ED&'<()'&#j#2%(IA#(2&#F.<**393&'#K%).5#F<.F).<(&#8H"&*I#m#8H%)(i"&*I#m#8H,%,&i"&*I#m#8H2%(i"&*I#<,5#8H,%I#m#8H%)(i,%I#m#8H,%,&i,%I#m#8H2%(i,%I#12&#&,5#'&*).(#K%).5#!&#(2&#F.<**393&'#*&.&F(3,/#(2&#E<J3E)E#6<.)&#%9#(2&#(K%#<,5#F2%%*&#&3(2&'#"&*#%'#,%;##Z<I6&#+<"&*#2<*#!&&,#)*&5#9%'#(<*L*#*)F2#<*#*%'(3,/#%)(#,&K*#<'(3F.&*#H]%<F23E*A#?RRCI;

Bayesian Networks

Z<I6&#+<"&*#<**)E&*#(2<(#<..#<(('3!)(&#6<.)&*#<'&#F%,53(3%,<.."#3,5&D&,5&,(#/36&,#<#(<'/&(#<(('3!)(&#6<.)&;##12)*A#(2&'&#,&&5*#(%#!&#<#K<"#(%#F.<**39"#*%E&#<(('3!)(&*#<*#F%,53(3%,<.."#3,5&D&,5&,(A#!)(#,%(#%(2&'*; 12&#*%.)(3%,#3*#<#+<"&*3<,#,&(K%'L;

Definition: *Attributes are **conditionally independent** of one another if given the value of one or more attributes $Y_1 \ldots Y_m$ determines the value of attributes $X_1 \ldots X_m$ independent of values of attributes $Z_1 \ldots Z_m$* H>3(F2&..A#?RRUI;

M,#(23*#K<"A#<#,&(K%'L#3*#F'&<(&5#!"#F%,,&F(3,/#<(('3!)(&*#(2<(#<'&#,%(#F%,53(3%,<.."#3,5&D&,5&,(#%9#%,&#<,%(2&'#<,5#F<.F).<(3,/#(2&3'#F%,53(3%,<.#D'%!<!3.3(";##12&'&9%'&#<#*"'*(&E#%9#,%5&*#3*#F'&<(&5#*3E3.<'#(%#(2&#*"'*(&E#F'&<(&5#!"#,&)'<.#,&(K%'L*;##M,#(23*#*3()<(3%,A#"%)#F<,#3,9&'#(<'/&(#6<.)&*#9%'#<(('3!)(&*#!"#F<.F).<(3,/#(2&#D'3%'#D'%!<!3.3("#%9#&<F2#6<.)&#/36&,#(2&#D<'&,(*#%9#(2<(#,%5&;##M,#(23*#K<"#"%)#F<,#F%,,&F(#,%5&*#(%/&(2&'#<,5#2<6&#<#,&(K%'L#%9#,%5&*#2<63,/#E<,"#D<'&,(*;##4,#&J<ED.&#%9#<#+<"&*3<,#,&(K%'L#3*#*2%K,#3,#V3/)'&*#[<#<,5#[!;

Figure 5a: Example of a Bayesian Network

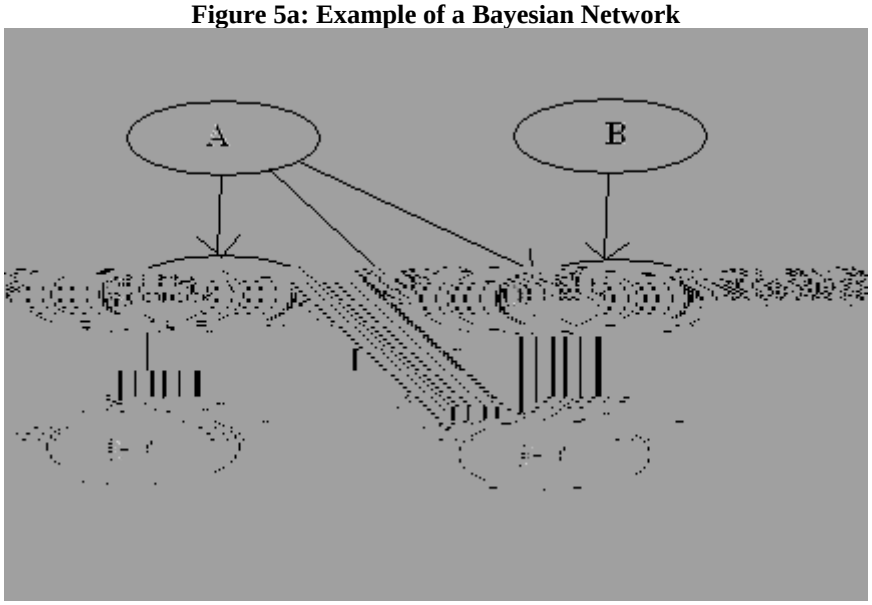


Figure 5a: Example of a Conditional Probability Table

	<A5	<A#n5	n<A5	n<A#n5
9	B;\	B;?	B;o	B;@
n9	B;C	B;R	B;@	B;o

M,#(23*#&J<ED.&#-#2<*#D<'&,(#`A#b#2<*#D<'&,(#-A#`#2<*#D<'&,(*#4#<,5#+A#<,5#VA#(2&#(<'/&(#<(('3!)(&A#2<*#D<'&,(*#4#<,5#`;##M(#3*#,%(#Q%3,&5#3,#.<"&'*#<*#3,#,&)'<.#,&(K%'L*A#<*#(2&#D<'&,(*#%9#,%5&*#F<,#<.*%#!&#(2&#D<'&,(*#%9#(2&3'#F23.5'&,;##12&'&9%'&#F<.F).<(3,/#(2&#D'%!<!3.3("#%9#V#!&3,/#9A#/36&,#(2&#5<(<#(2<(#4#3*#<#<,5#`#3*#5#K%).5#!&A#9%'#&J<ED.&A#

K%).5#!&#8HVj9i4 j<A` j5I;##Y*3,/#(2&#F%,53(3%,<.#D'%!<!3.3("#(<!.&#<!%6&A#(23*#6<.)&#
K%).5#(2&'&9%'&#!&#B;\;
Y,.3L&#,&)'<.#,&(K%'L*A#3,#K23F2#&''%'*#K<,(#(%#!&#E3,3E3=&5A#+<"&*3<,#,&(K%'L*
F<,#!&#('<3,&5#(%#E<J3E3=&#(2&#D'%!<!3.3("#%9#(2&#%!*&'6&5#5<(<#/36&,#(2&#,&(K%'L#
D<'<E&(&'*;# h &3/2(*#H/T f<ETB.15T f<ETW2<*&I%5Q#*239(<K%\)D5<1(F<) j9i4 j<A` j5#(2&#,&(K%'L#23*#6<.)

Machine Learning Work with Recommender Systems

>< , "#F%..<!%'<(36&#93.(&'3 , /#(&F2 , 3G)&*#2<6&#!&& , #)*&5#!&9%'&;##V%'#&J<ED.&A# $3 , /%#K<*#<#*%F3<.#3 , 9%'E<(3% , #93.(&'#)*&5#(%#(' '#< , 5#E<L&#E)*3F#'&F%EE& , 5<(3% , *#(%# .3*(& , &'*#H : 2<'5< , < , 5#S# > <&*A#?RR[I;## M , #*)F2A#'&F%EE& , 5<(3% , *#<'&#D<**&5#(2'%) / 2# pK%'5#%9#E%)(2qA#'<(2&'#(2< , #F% , (& , (A#*%#E)*3F#'&F%EE& , 5<(3% , *#E<'' #!&#5'<*(3F<..'' # 5399&'& , (#(2< , #(2%*&#(2<(#<#)*&'#2<*#.3*(& , &5#(%#!&9%'&;#12&'&#2<6&#!&& , #E< , ''#%(2&'#(&*(*#

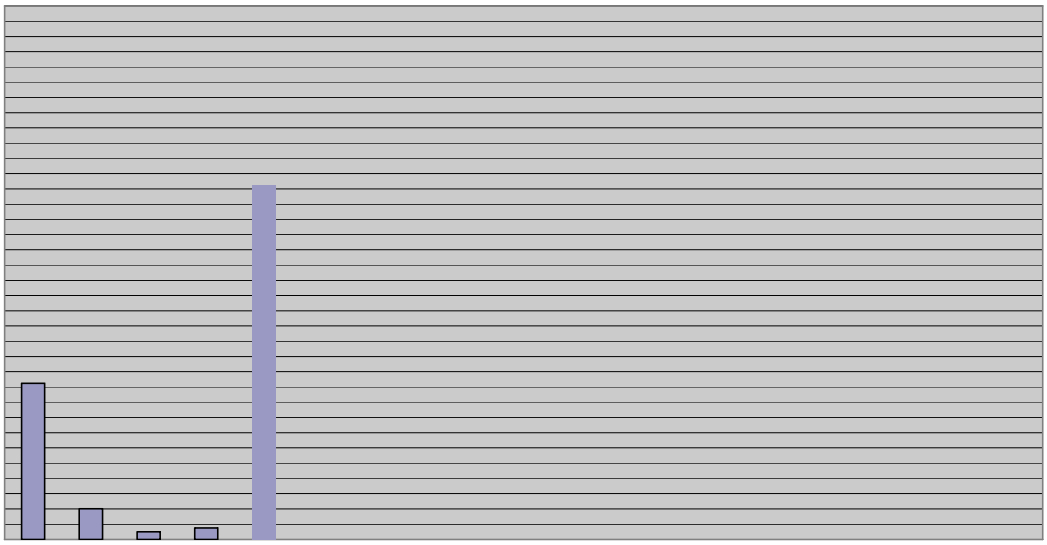
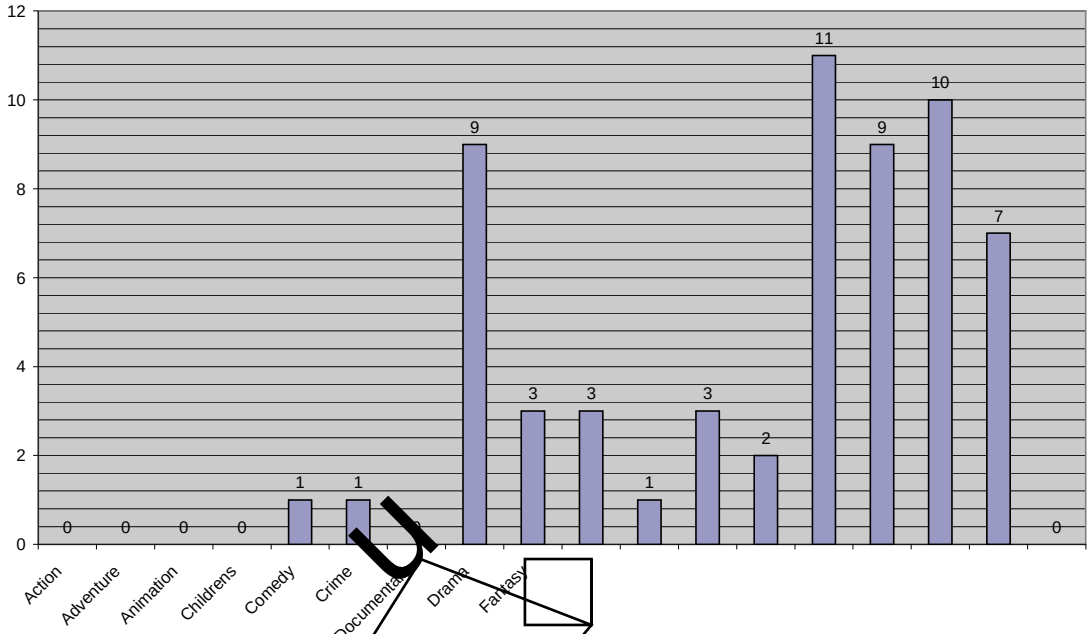
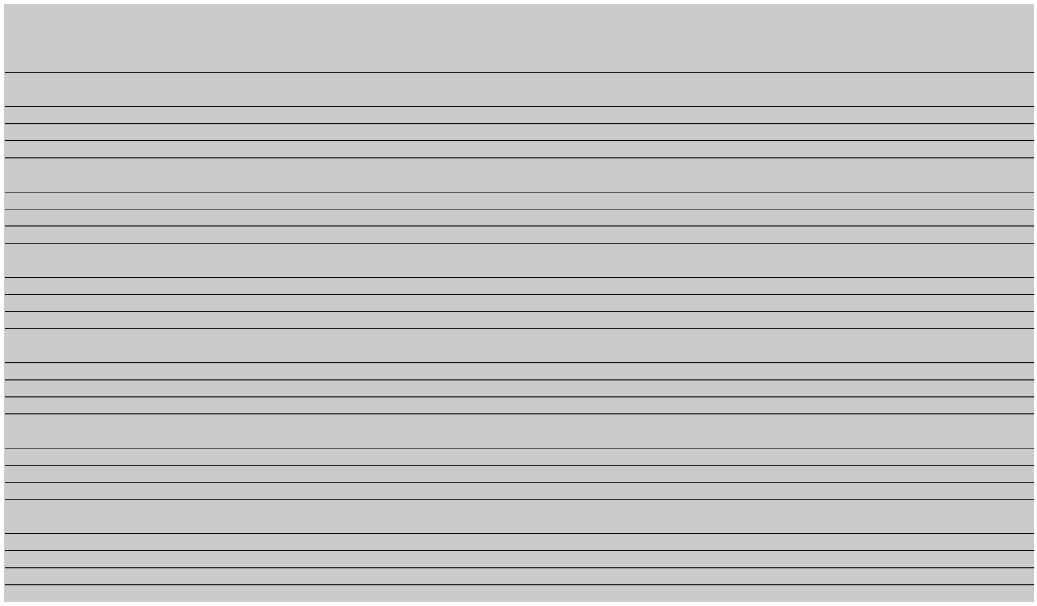


Figure 7c: Genre3 Distribution





Performance Metrics and Evaluation Protocol

M#5&F35&5#(%#&6<.)<(&#(2&#'&*).(*#%9#(&*(*#%9#(2&#E<F23,&#.&<',3,/#<./%'3(2E*#%,#(2&#5<(<*&(#!"#)*3,/#<FF)'<F"A#D'&F3*3%,A#<,5#(%DNZ#D'&F3*3%,;

Definition: ***Accuracy*** *is the percentage of instances that are correctly classified by the system.*

Definition: ***Precision*** *is the percentage of like predictions that agree with the user's taste.*

-&'(<3,#E<F23,&#.&<',3,/#<./%'3(2E*A#*)F2#<*#+<"&*3<,#<,5#,&)'<.#,&(K%'L*A#

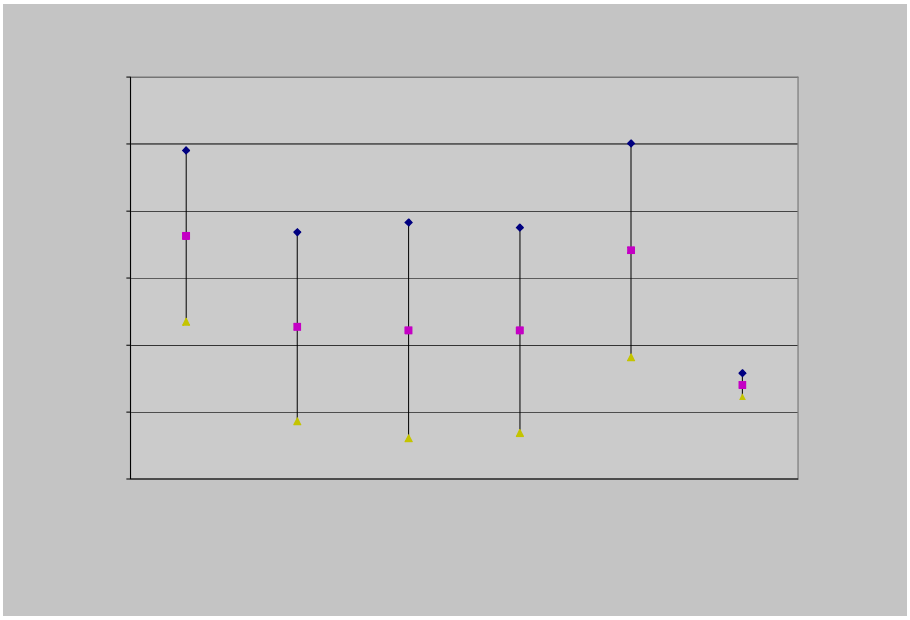
Results

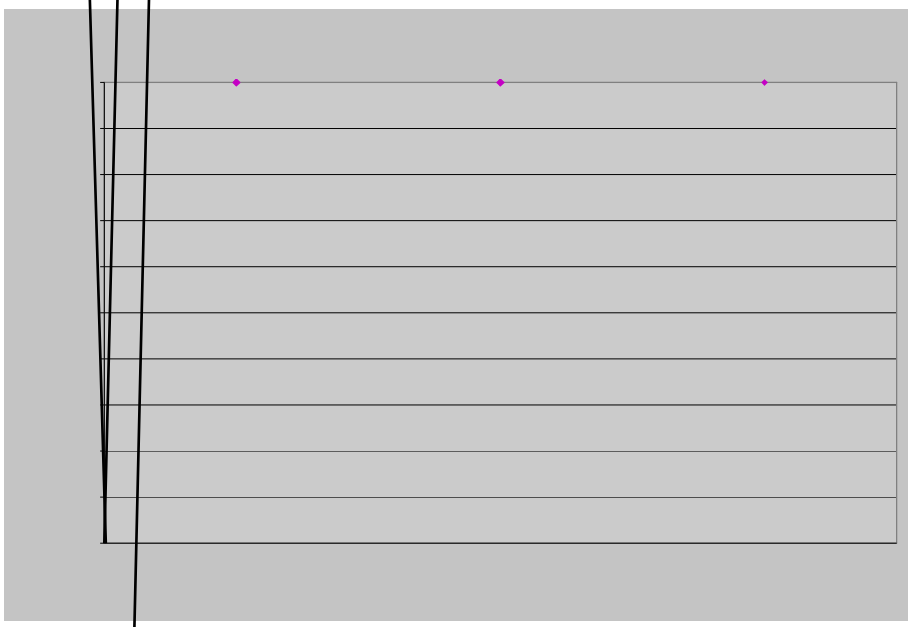
Attribute Selection

M#(&*(&5#!%(2#D'3,F3D<.#F%ED%,&,(*#<,5#+&*(V3'*(#(%#*&.&F(#(2&#<(('3!)(&*#(%#!&#9&5#<*#3,D)(*#(%#(2&#6<'3%)*#E<F23,&#.&<',3,/#<./%'3(2E*;##

Principal Components

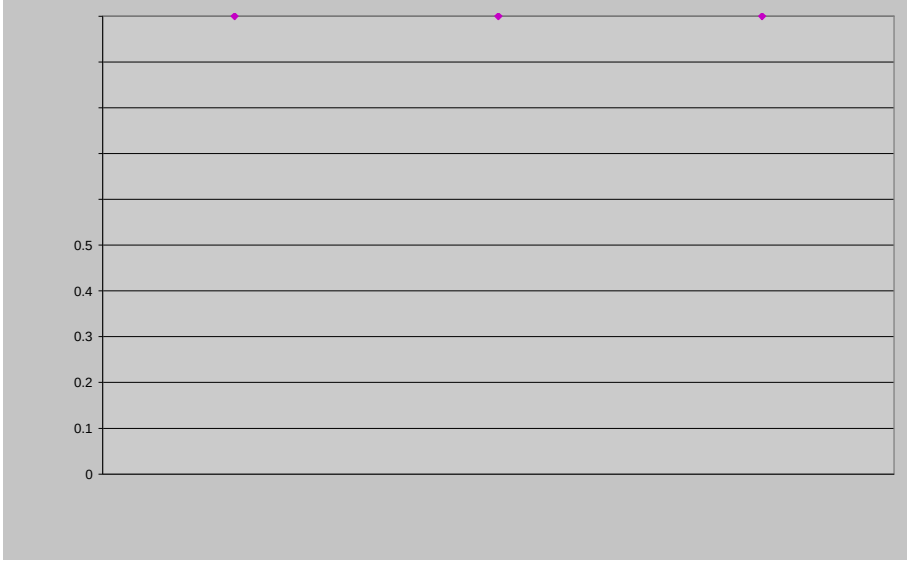
V3/)'&*#o<A#o!A#<,5#oF#*2%K#(2&#D&'9%'E<,F&#%9#(2&#5399&'&,(#E<F23,&#.&<',3,/#<./%'3(2E*#%6&'#5<(<#(2<(#K<*#D'&D'%F&**&5#)*3,/#D'3,F3D<.#F%ED%,&,(*#<(('3!)(&#*&.&F(3%,;

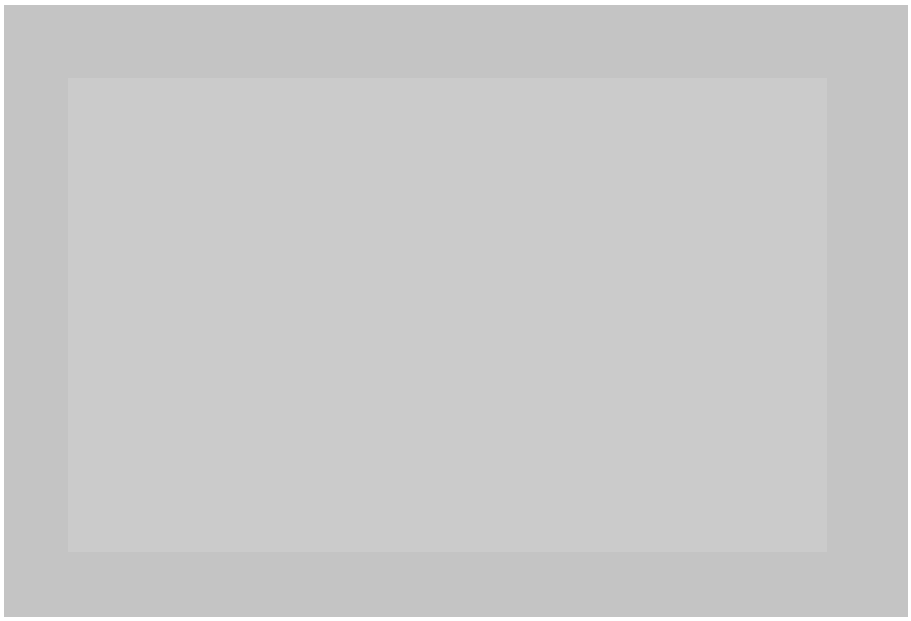




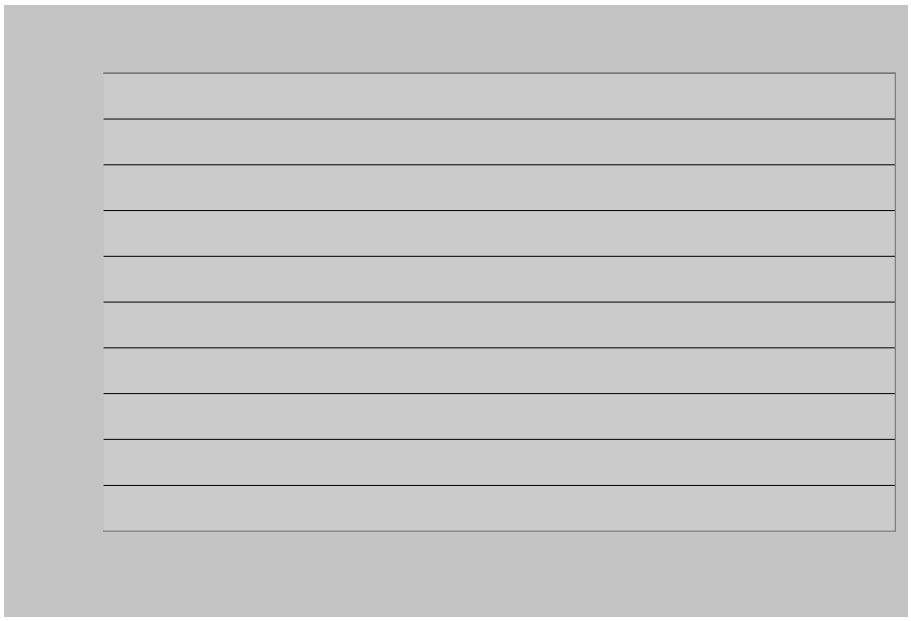
BestFirst

Figure 9c: Top-10 Precision Using BestFirst Attribute Selection





Neural Networks



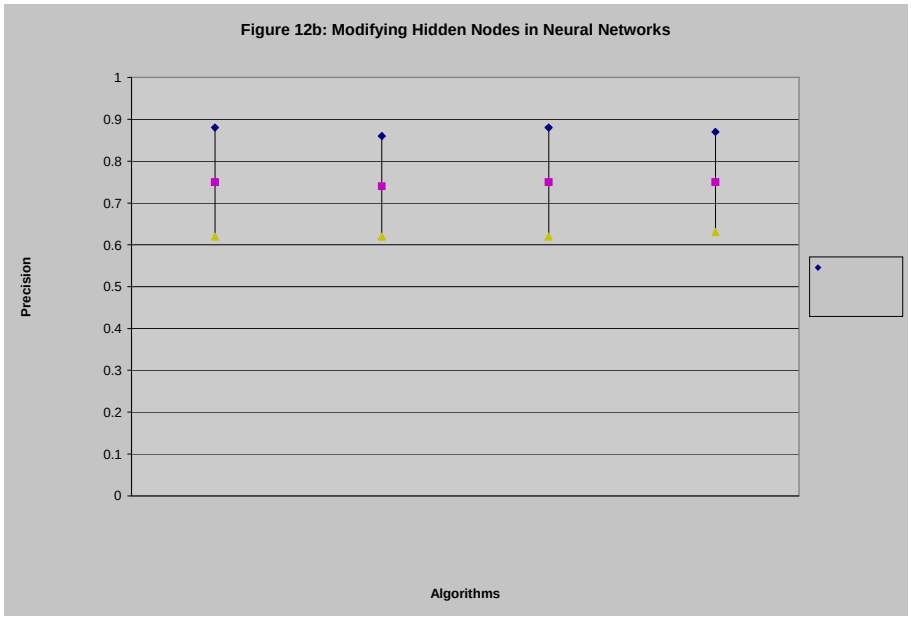
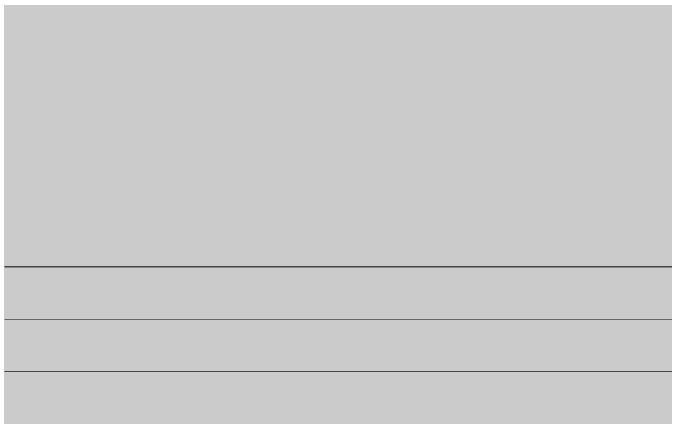
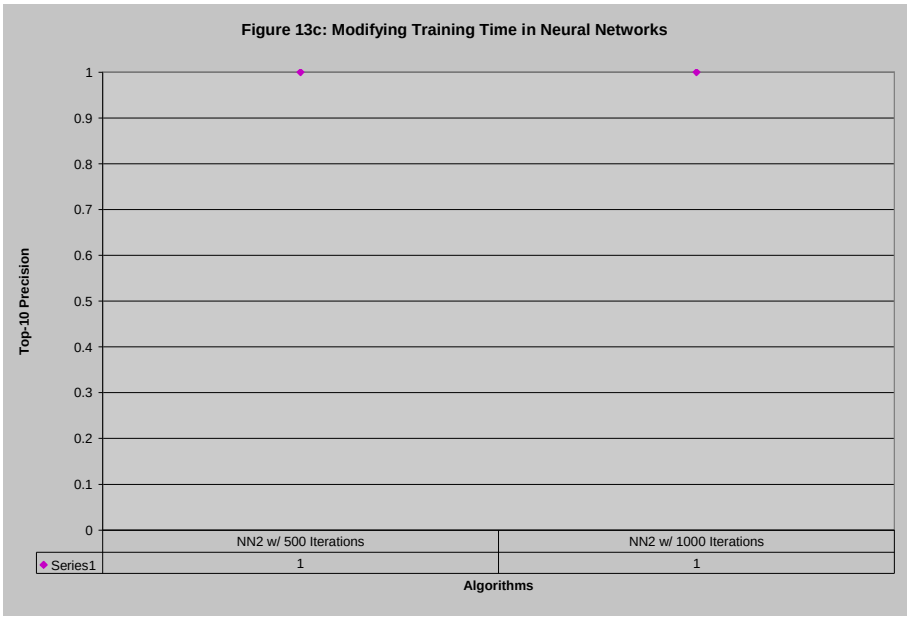


Figure 12b: Modifying Hidden Nodes in Neural Networks
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Precision
Algorithms

Number of Training Epochs

M#(2&,#5&F35&5#(%#3,F'&<*&#(2&#*(<,5<'5#('<3,3,/#(3E&#%,#(2&#5<(<#9'%E#[BB#3(&'<(3%,*#(%#?BBB#3(&'<(3%,*;##12&#'&*).(*#<'&#*2%K,#3,#V3/)'&*#?^<A ?^!#<,5#?^F;



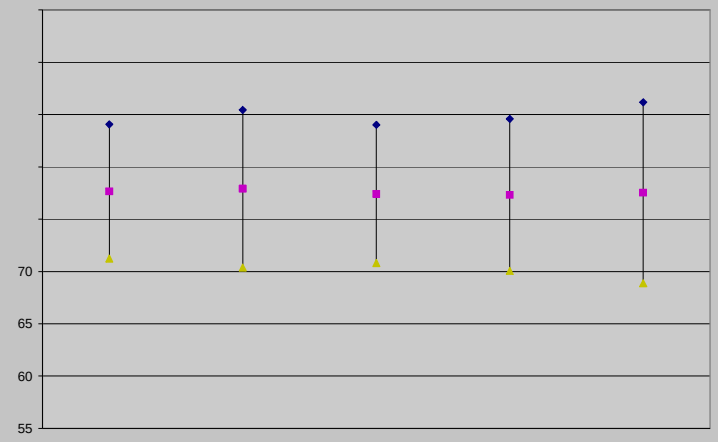


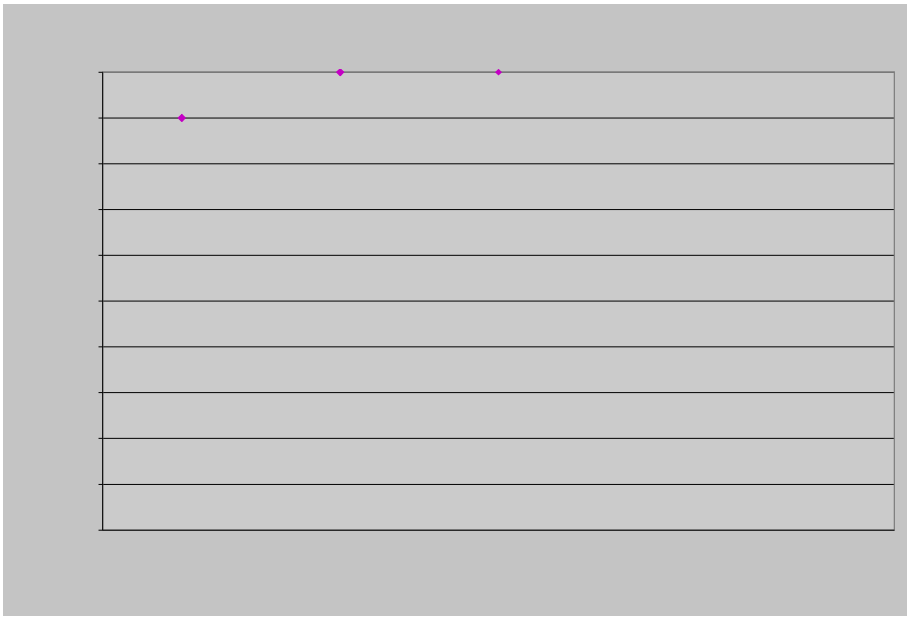
X%%L3,/#<(#(2&#'&*).(*A#(%DN?B#D'&F3*3%,#K<*#(2&#*<E&#5&*D3(&#3,F'&<*3,/#(2&#('<3,3,/#(3E&;#4FF)'<F"#*.3/2(."#5&F'&<*&5A#<,5#(2&#*(<,5<'5#5&63<(3%,#*.3/2(."#3,F'&<*&5;##8'&F3*3%,#<.*%#*.3/2(."#5&F'&<*&5;##12&'&9%'&A#5&*D3(&#3,F'&<*3,/#(2&#('<3,3,/#(3E&A#(2&#<FF)'<F"#<,5#*(<,5<'5#5&63<(3%,#535#,%(#2<6&#<#*3/,393F<,(#F2<,/&;##12&'&9%'&A#M#535#,%(#!%(2&'#3,F'&<*3,/#(2&#('<3,3,/#(3E&#9%'#9)()'&#&JD&'3E&,(*;

Learning Rate

M#<.*%#F%,*35&'&5#E%539"3,/#(2&#.&<',3,/#'<(&#K3(23,#,&)'<.#,&(K%'L*#(%#5&(&'E3,&#K2&(2&'F%,6&'/&,F&(%#<#!&((&'#E3,3E)E#F%).5#!&#<F23&6&5;##12&#'&*).(*#<'&#53*D.<"&53,#V3/)'&*?\<A?\!#<,5#?\F;

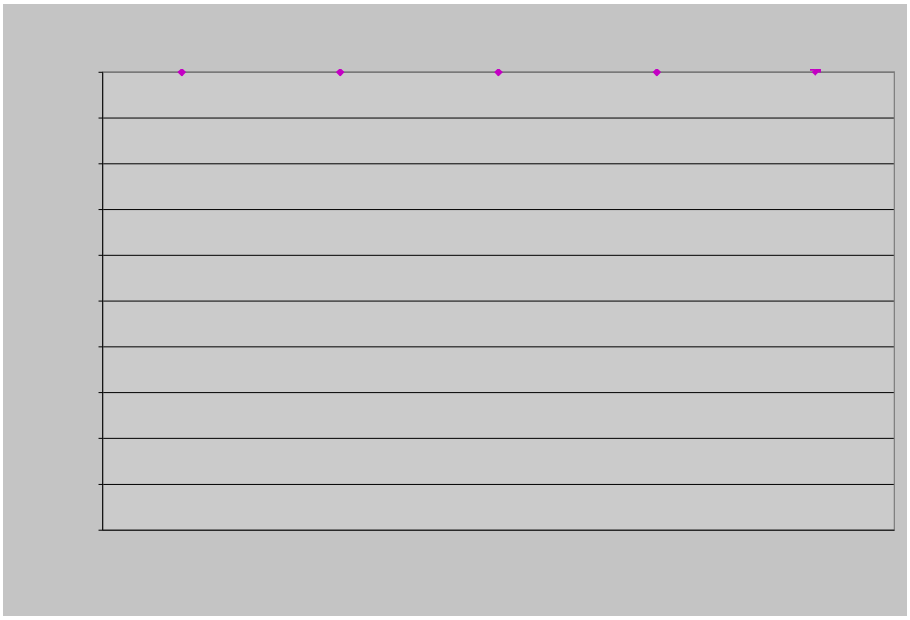
Figure 14a: Modifying the Learning Rate in Neural Networks





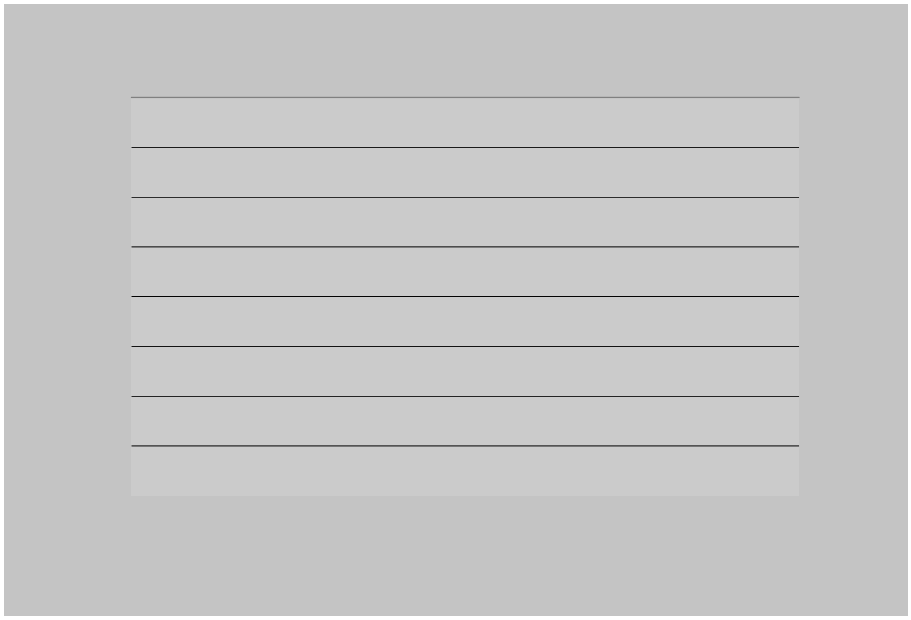
Bayesian Networks

M#5&F35&5#(%#(&*(#(2&#E<J3E)E#,)E!&'#%9#D<'&,(*#&<F2#,%5&#F<,#2<6&#%,#(2&#



Collaborative Data Only

V3/) '&*#?C<A#?C! <, 5#?CF#*2%K#(2&#/ '<D2*#9%'#F%..<!%'<(36&N%, ."#+<"&*3<, #, &(K%'L*;



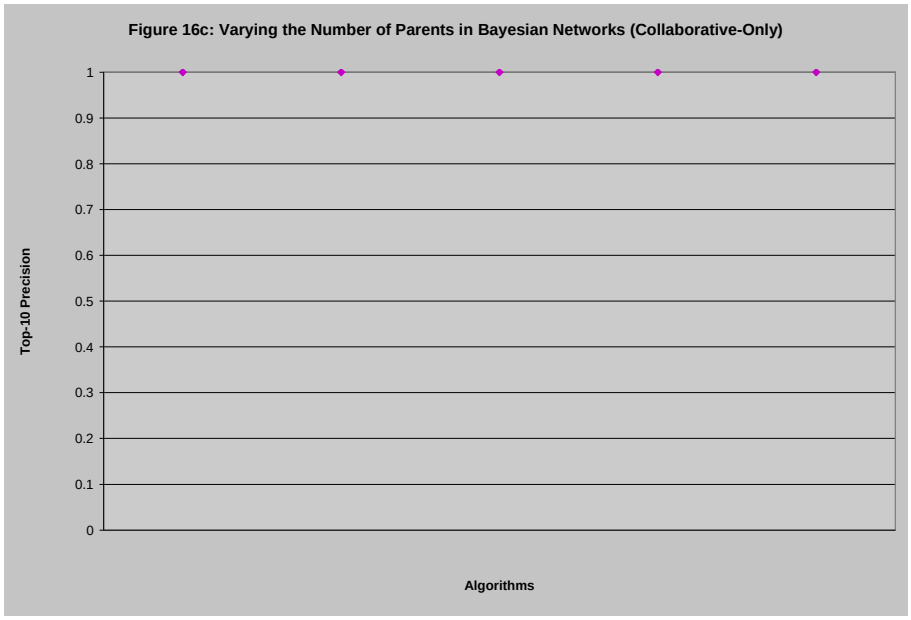
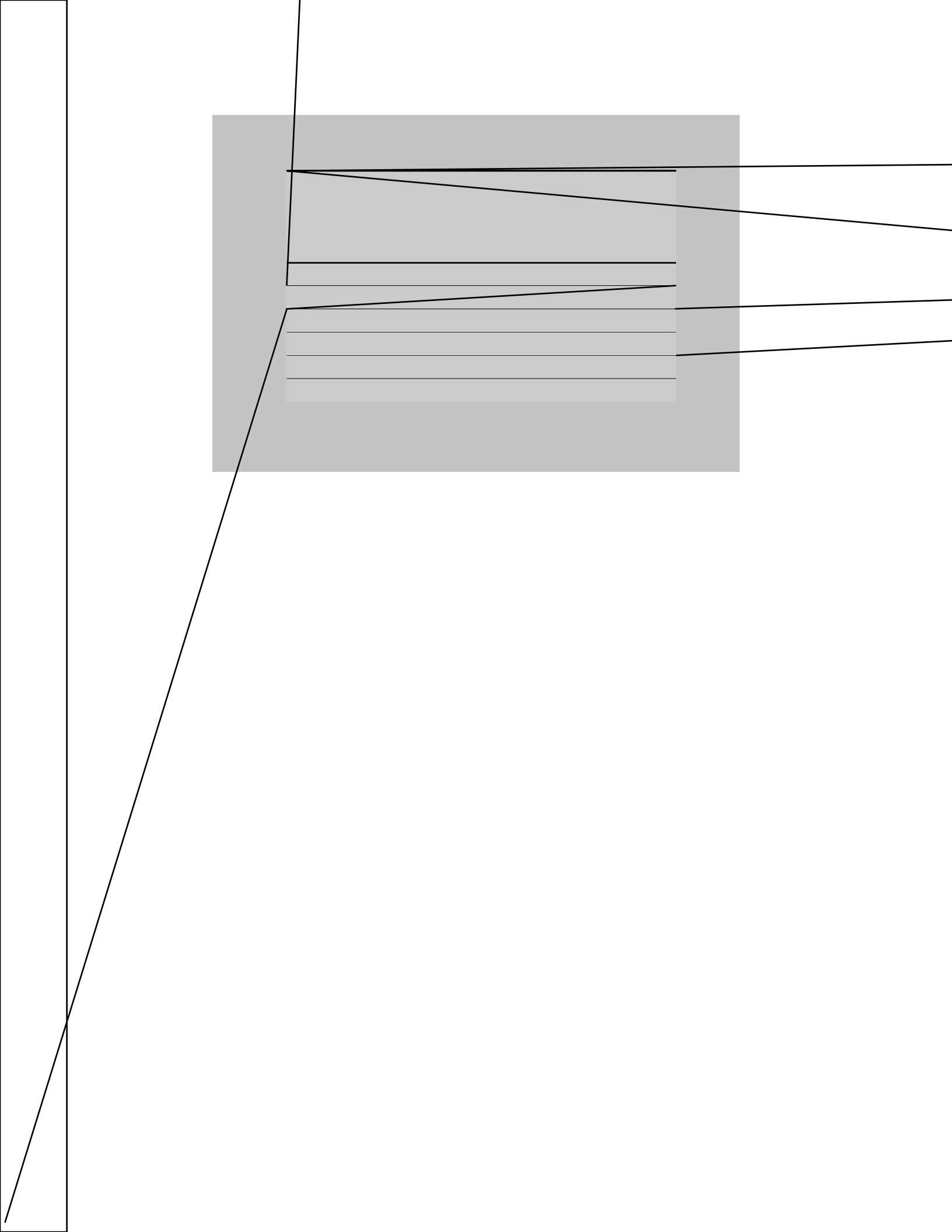
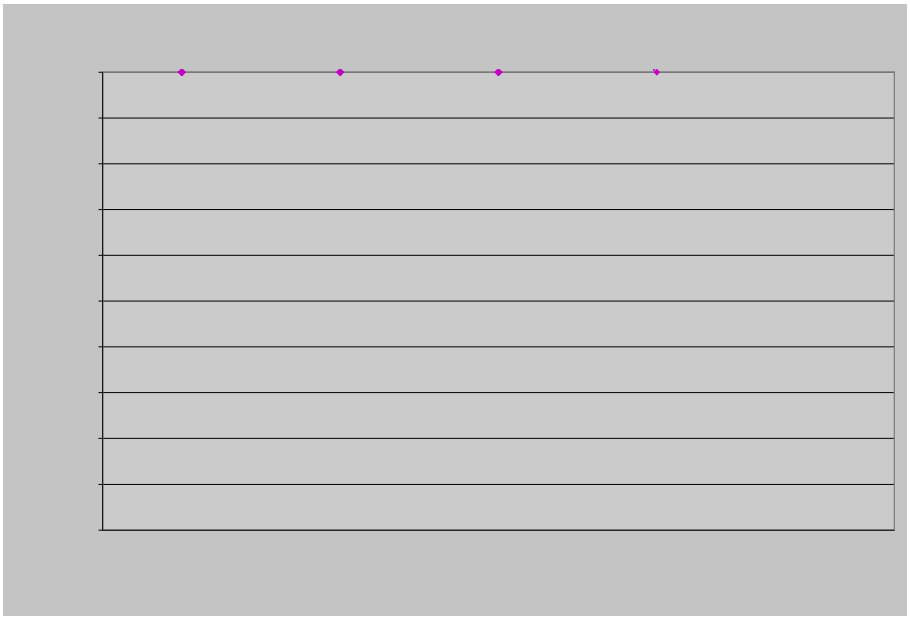


Figure 16c: Varying the Number of Parents in Bayesian Networks (Collaborative-Only)
Top-10 Precision
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Algorithms



Enhanced Content Data



Conclusions and Future Work

12&#D)'D%*&#%9#(23*#(2&*3*#K<*#(%#<,<."=&#E<F23,&#.&<',3,/#(&F2,3G)&*#<,5#<DD."#(2&E#(%#'&F%EE&,5&'#*"'*(&E*#3,#%'5&'#(%#E<L&#'&F%EE&,5<(3%,*#9%'#E%63& D<('%,*;##+"#<DD."3,/#5399&'&,(#*('<(&/3&*#(%#!%(2#(2&#5<(<*&(#<,5#(2&#5399&'&,(#<./%'3(2E*A#(2&#2%D&#K<*#(%#93,5#<,#%D(3E3=&5#K<"#%9#D'&53F(3,/#K2&(2&'#)*&'*#K%).5#.3L&#F&'(<3,#E%63&*#!<*&5#%,#F%,(&,(#3,9%'E<(3%,#<!%)(#(2&#E%63&#<,5#F%..<!%'<(36&#3,9%'E<(3%,#9'%E#%(2&'#E%63&/%&'*0#'<(3,/*;##V'%E#(2&#K%'L#5%,&A#3(#F<,#!&#F%,F.)5&5#(2<(#F&'(<3,#E<F23,&#.&<',3,/#<./%'3(2E*#F<,#3,5&&5#D&'9%'E#K&..#3,#E<L3,/#E%63&#'&F%EE&,5<(3%,*;

M,#(&'E*#%9#(%DNZ#D'&F3*3%,A#,&)'<.#,&(K%'L*#<,5#D'3,F3D<.#F%ED%,&,(*#D&'9%'E#&J('&E&."#K&..#9%'#<#.%K#6<.)&#%9#Z#H,)E!&'#%9#E%63&*#'&F%EE&,5&5I;##123*#3*#6&'"#3ED%'(<,(#9%'#E%63&#'&F%EE&,5<(3%,*A#<*#E%*(#)*&'*#K%).5#%,."#K<,(#(%#*%'(#(2'%)/2#?B#%'#9&K&'#E%63&*#3,#%'5&'#(%#D3FL#%,&;##12&#9<F(#(2<(#(%DNZ#D'&F3*3%,#'&E<3,*#23/2#)D#(%#Z j @BB#3*#<.*%#<#/%%5#*%)'F&#%9#%(2&'#*%'(#%9#'&F%EE&,5<(3%,*A#*)F2#<*#'&F%EE&,5<(3%,*#9%'#D&%D.&#K2%#!)"#%'#'&,(#`T`*#%'#K<(F2#E%63&*#%,#(&.&63*3%,;##12%*&#D&%D.&#E3/2(#,&&5#<#!3//&'#.3*(#<*#(2&"#K<,(#<#.3*(#%9#E%63&*#(%#K<(F2#'<(2&'#(2<,#Q)*(#<#*&.&F(#9&K;

+<"&*3<,#,&(K%'L*#(<L&#E)F2#.&**#('<3,3,/#(3E&#(2<,#,&)'<.#,&(K%'L*#K23.&#2<63,/#(2&#*<E&#(%DN?B#D'&F3*3%,A#*%#3(#F%).5#!&#E%'&#)*&9).#9%'#'&N('<3,3,/#K2&,#)*&'*#<55#(2&3'#.3L&*#<,5#53*.3L&*#3,(%#<#*"'*(&E;##4.*%A#3,#(&'E*#%9#<FF)'<F"A#+<"&*3<,#,&(K%'L*#53*D.<"#<#*3/,393F<,(#3ED'%6&E&,(#%6&'#,&)'<.#,&(K%'L*A#K3(2#!%(2#(2&#*(<,5<'5#6&'*3%,#<,5#E%5393&5#6&'*3%,#%9#&''%'#D'%D</<(3%,A#<*#K&..#<*#5&F3*3%,#('&&*A#Z<I6&#+<"&*A#?$#<,5#B$#F.<**393&'*;##+<"&*3<,#,&(K%'L*#D&'9%'E#K&..#K3(2#!%(2#D'3,F3D<.#F%ED%,&,(*#<,5#(2&#+&*(V3'*(#<((('3!)(&#*&.&F(3%,A#*%E&(23,/#(2<(#,%,&#%9#(2&#%(2&'#<./%'3(2E*#(&*(&5#5%;##M(#3*#F.%*&#K3(2#?$#3,#D'3,F3D<.#F%ED%,&,(*#!)(#(23*#F%).5#!&#5)&#(%#(2&#*3=&#%9#(2&#5<(<*&(;##4#5<(<*&(#K3(2#E%6&#E%63&*#9%'#&<F2#"&<'#F%).5#D%**3!."#.%K&'#?$0*#D&'9%'E<,F&;##

+&*(V3'*(#<DD&<'*#(%#!&#<,#&99&F(36&#(&F2,3G)&#9%'#<(('3!)(&#*&.&F(3%,;##M('&5)F&*#(2&#,)E!&'#%9#<(('3!)(&*#F%,*35&'<!."A#/36&*#?BB r #(%DN?B#D'&F3*3%,#<,5#3,F'&<*&*#(2&#E<J3E)E#<FF)'<F"#9'%E#(2&#B$#!<*&.3,&;##+&*(V3'*(#3*#<.*%#9<*(&'#(2<,#D'3,F3D<.#F%ED%,&,(*#3,#(2<(#3(#*&.&F(*#9&<()'&*#'<(2&'#(2<,#<DD."3,/#K&3/2(*#(%#<..#%9#(2&EA#*%#3(#3*#!&((&'#3,#(&'E*#%9#<FF)'<F"A#D'&F3*3%,#<,5#(3E&A#!)(#,%(#3,#%6&'<..#(%DNZ#D'&F3*3%,;

12&'&#5%&*#,%(#<DD&<'#(%#!&#<#!3/#F2<,/&#3,(%DN?B#D'&F3*3%,A#<FF)'<F"#%'#D'&F3*3%,#K2&,#F%ED<'3,/#,&)'<.#,&(K%'L*#K3(2#(2&#*(<,5<'5#6&'*3%,#%9#&''%'#!<FLD'%D</<(3%,#<,5#(2&#E%5393&5#6&'*3%,;##>%'&#(&*(3,/#K%).5#,&&5#(%#!&#5%,&#%6&'#.%,/&'#<E%),(*#%9#(3E&#(%#*&&#39#%,&#3*#*3/,393F<,(."#!&((&'#(2<,#(2&#%(2&';

>%'&#(&*(3,/#F%).5#!&#5%,&#%,#(2&#D<'<E&(&'#*&((3,/*#%9#,&)'<.#,&(K%'L*#3,#%'5&'#(%#5&(&'E3,&#K2<(#D<'<E&(&'#6<.)&*#%D(3E3=&#<FF)'<F"A D'&F3*3%,#<,5#(%DN?B#D'&F3*3%,;##M(#<DD&<'*#(2<(#<5Q)*(3,/#(2&#('<3,3,/#(3E&A#.&<',3,/#'<(&#<,5#%(2&'#6<'3<!.&*#5%&*#,%(#*3/,393F<,(."#F2<,/&#(2&#<FF)'<F"#<,5#(2&#D'&F3*3%,#%9#(2&#'&F%EE&,5<(3%,*;

12&#'&*).(*#%9#(2&#%D(3E<.#<./%'3(2E*#&JF&&5#(2&#!&,F2E<'L*#*&(#!"#B$*3/,393F<,(.";##12&#!&*(#D&'9%'E3,/#<./%'3(2E*#2<6&#?BB r #(%DN?B#D'&F3*3%,A#K23F2#3*#<#6&'"#3ED%'(<,(#*(<(3*(3F;#12&"#<.*%#3,F'&<*&5#(2&#<FF)'<F"#<,5#D'&F3*3%,#9'%E#C@ r #(%#o^ r <,5#C@ r #(%#oB r A#'&*D&F(36&.";

M(#<DD&<'*#(2<(#)*3,/#F%,(&,(N%,."A#F%..<!%'<(36&N%,."A#%'#<#F%E!3,&5#5<(<*&(#5%&*#,%(#F2<,/&#(2&#(%DN?B#D'&F3*3%,A#<*#<..#(2'&&#5<(<*&(#6&'*3%,*#/<6&#?BB r #(%DN?B#D'&F3*3%,;##?BB r #(%DN?B#D'&F3*3%,#3,#F%,(&,(N%,."#5<(<#3*#&*D&F3<.."#/%%5#9%'#<#,&K#E%63&#9%'#K23F2#,%#F%..<!%'<(36&#5<(<#3*#<6<3.<!.&;##-%..<!%'<(36&#5<(<#%,."#D'%6&5#E%'&#)*&9).#3,#

3,F'&<*3,/#%6&'<..#<FF)'<F";## 0%K&6&'A#3(#K%).5#!&#3,(&'&*(3,/#(%#*&&#K2<(#K%).5#2<DD&,#39#&6&,#E%'&#F%,(&,(#5<(<#K<*#<55&5A#*)F2#<*#<F(%'*A#<*#*%E&#E%63&/%&'*#2<6&#9<6%'3(&#<F(%'*#<,5#<'&#E%'&#.3L&."#(%#*&&#93.E*#!"#(2%*&#<F(%'*;

V3,<.."A#+<"&*3<,#<,5#,&)'<.#,&(K%'L#(&F2,3G)&* *2%).5#!&#(&*(&5#%,#%(2&'#5<(<*&(*#(2<(#F%,(<3,#E%'&#5<(<;##12&#5<(<*&(#F2%*&,#K<*#'<(2&'#*E<..#5)&#(%#(2&#(3E&#F%ED.&J3(3&*#%9#<#.<'/&'#5<(<*&(;##M(#*2%).5#!&#5&(&'E3,&5#K2&(2&'#(2&#F%,F.)*3%,#(<L&,#9'%E#(23*#*&(#%9#&JD&'3E&,(*#<DD.3&*#(%#.<'/&'#5<(<*&(*;##4553(3%,<.#(<'/&(#)*&'*#*2%).5#!&F%,*35&'&5 <*#K&..;

References

4.6<'&=A#:; 4;A#$)3=A#-;A#_<K<(%A#1;A#S#_%/&.A#h;#H@BBCI;#Z&)'<.#bJD&'(#Z&(K%'L*#9%'#V<*(&'#-%E!3,&5#-%..<!%'<(36&#<,5#-%,(&,(N+<*&5#$&F%EE&,5<(3%,;#M,#*Journal of Computational Methods in Sciences and Engineering*,#(%#<DD&<';

+<.<!<,%63FA#>;A#S#:2%2<EA#W;#H?RRUI;#-%E!3,3,/#F%,(&,(N+<*&5#<,5#F%..<!%'<(36&#'&F%EE&,5<(3%,;#M,#*Communications of the ACM*A#\BH^IA#DD;#CCNU@;

+3..*)*A#`;A#S#8<==<,3A#>;];#H?RRoI;#X&<',3,/#F%..<!%'<(36&#3,9%'E<(3%,#93.(&'*;#M,#*Proceedings of the Fifteenth International Conference on Machine Learning*A#DD;#\CN[\;

+3*2%DA#-;#>;#H?RRCI;#*Neural networks for pattern recognition*;#PJ9%'5A#b,/.<,57#PJ9%'5#Y,36&'*3("#8'&**;

+'&&*&A#];#:;A#0&FL&'E<,A#`;A#S#_<53&A#-;#H?RRoI;#bED3'3F<.#<,<."*3*#%9#D'&53F(36&#<./%'3(2E*#9%'#F%..<!%'<(36&#93.(&'3,/;#M,#*Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*A#DD;#\^g[@;

-%%D&'A#c;A#S#0&'*L%63(*A#b;#H?RR@I;#4#+<"&*3<,#E&(2%5#9%'#(2&#3,5)F(3%,#%9#D'%!<!3.3*(3F#,&(K%'L*#9'%E#5<(<;#M,#*Machine Learning*A#RA#DD;#^BRN^\U;

0<'D&'A#V;A#X3A#f;A#-2&,AW;#S#_%,*(<,A#];#H@BB[I;#4,#bF%,%E3F#>%5&.#%9#Y*&'#$<(3,/#3,#<,#P,.3,&#$&F%EE&,5&'#:"*(&E;#M,#*Proceedings of the 10th International Conference on User Modeling*;

]%<F23E*A#1;#H?RRCI;#*A probabilistic analysis of the Rocchio algorithm with TFIDF for text Categorization*A#H-%ED)(&'#:F3&,F&#1&F2,3F<.#$&D%'(#->YN-:NRCN??oI;#-<',&/3&#>&..%,#Y,36&'*3(";

>3,*L"A#>;A#S#8<D&'(A#:;#H?RCRI;#*Perceptrons*;#-<E!'35/&A#>47#>Mb#8'&**;

>3(F2&..A#1;#>;#*Machine Learning*;#H?RRUI;+%*(%,A#>47#12&#>Fc'<KN03..#-%ED<,3&*A#M,F;

a)3,.<,A#];#$;#H?RoCI;#M,5)F(3%,#%9#5&F3*3%,#('&&*;#M,#*Machine Learning*A#?H?IA#DD;#o?N?BC;

a)3,.<,A#];#$;#H?RR^I;#*C4.5: Programs for Machine Learning*;#:<,#><(&%A#-47#>%'/<,#
_<)9E<,,;

Appendix

Top-N Precision w/ NN1 Code

```
3ED%'(#Q<6<;3%;+)99&'&5$&<5&'u
3ED%'(#Q<6<;3%;+)99&'&5 h '3(&'u
3ED%'(#Q<6<;3%;V3.&$&<5&'u
3ED%'(#Q<6<;3%;V3.& h '3(&'u

3ED%'(#K&L<;F%'&;M,*(<,F&u
3ED%'(#K&L<;F%'&;M,*(<,F&*u
3ED%'(#K&L<;F.<**393&'*;9),F(3%,*;>).(3.<"&'8&'F&D('%,u
3ED%'(#K&L<;F.<**393&'*;b6<.)<(3%,u


D)!.3F#F.<**#1%DZ8'&F3*3%,#v

        D)!.3F#*(<(3F#6%35#E<3,H:('3,/wx#<'/*Iv

                ('"v
                        M,*(<,F&*#5<(<#j#,&K#M,*(<,F&*H
                                        ,&K#+)99&'&5$&<5&'H
                                                ##,&K#V3.&$&<5&'HyF%E!+V;<'99yIIIu
                        kk*&((3,/#F.<**#<(('3!)(&
                        5<(<;*&(-.<**M,5&JH5<(<;,)E4(('3!)(&*HI#N#?Iu


                        kk#.%<5#),.<!&.&5#5<(<

                        >).(3.<"&'8&'F&D('%,#E.D#j#,&K#>).(3.<"&'8&'F&D('%,HIu
                        E.D;*&(0355&,X<"&'*Hy^yIu              #kk#*&(#(2&#2355&,#,%5&*
                        E.D;*&(X&<',3,/$<(&HB;@Iu #kk#*&(#(2&#.&<',3,/#'<(&
                        E.D;*&(1'<3,3,/13E&H[BBIu              #kk#*&(#(2&#('<3,3,/#(3E&
                        E.D;!)3.5-.<**393&'H5<(<Iu###kk#!)3.5#F.<**393&'

                        M,*(<,F&*#),.<!&.&5#j#,&K#M,*(<,F&*H
                                                                        ,&K#+)99&'&5$&<5&'H
                                                                        ##,&K#
V3.&$&<5&'HyF%E!+V;<'99yIIIu
                        kk#*&(#F.<**#<(('3!)(&
                        ),.<!&.&5;*&(-.<**M,5&JH),.<!&.&5;,)E4(('3!)(&*HI#N#?Iu
                        kk#F'&<(&#F%D"
                        M,*(<,F&*#.<!&.&5#j#,&K#M,*(<,F&*H),.<!&.&5Iu
                        kk#.<!&.#3,*(<,F&*
                        M,*(<,F&wx#3,*(<,F&*#j#,&K#M,*(<,F&w?Bxu
                        5%)!.&wx#6<.)&*#j#,&K#5%)!.&w?Bxu

                        9%'H3,(#3#j#Bu#3#z#?Bu#3eeIv
                                3,*(<,F&*w3x#j#,)..u
                                6<.)&*w3x#j#Bu
                        {

                        kk'<,L#3,*(<,F&*#N#%,."#L&&D#(%D#Z
                        9%'#H3,(#3#j#Bu#3#z#),.<!&.&5;,)EM,*(<,F&*HIu#3eeI#v
                        ##5%)!.&#wxF.*X<!&.#j#E.D;53*('3!)(3%,V%'M,*(<,F&H),.<!&.&5;3,*(<,F&H3IIu
```

```
                    ##39H6<.)&*wBx#z j #F.*X<!&.w?xI#v#kk.3L&#D'%!<!3.3("
                    ##          9%'H3,(#Q# j #?u#Q#z#?Bu#QeeIv
                                39H6<.)&*wQx#z j #F.*X<!&.w?x#SS#Q#| j #RIv
                                        6<.)&*wQN?x# j 6<.)&*wQxu
                                        3,*(<,F&*wQN?x# j #3,*(<,F&*wQxu
                                {
                                &.*&#39#HQ# j j #RIv
                                        6<.)&*wQN?x# j 6<.)&*wQxu
                                        3,*(<,F&*wQN?x# j #3,*(<,F&*wQxu
                                        6<.)&*wQx# j #F.*X<!&.w?xu
                                        3,*(<,F&*wQx# j #),.<!&.&5;3,*(<,F&H3Iu
                                {
                                &.*&v
                                        6<.)&*wQN?x# j #F.*X<!&.w?xu
                                        3,*(<,F&*wQN?x# j #),.<!&.&5;3,*(<,F&H3Iu
                                        !'&<Lu
                                {
                        {
                    ##{
                    {

                    .<!&.&5;5&.&(&HIu
                    9%'H3,(#3# j #Bu#3#z#?Bu#3eeI
                            39H3,*(<,F&*w3x#| j #,)..I
                                    .<!&.&5;<55H3,*(<,F&*w3xIu

                    kk&6<.)<(&#(2&#3,*(<,F&*
                    b6<.)<(3%,#&6<.# j #,&K#b6<.)<(3%,H.<!&.&5Iu
                    &6<.;F'%**T<.35<(&>%5&.H
                    ######E.DA#.<!&.&5A#?BA#.<!&.&5;/&($<,5%EZ)E!&'c&,&'<(%'H?IIu
                    :"*(&E;%)(;D'3,(.,H&6<.;(%:)EE<'":('3,/H(')&IIu

                    kkK'3(&#3,*(<,F&*#H(%#5%)!.&#F2&FL#,%#5%)!.&*I
                    +)99&'&5 h '3(&'#K'3(&'# j #,&K#+)99&'&5 h '3(&'H
                    ########################,&K#V3.& h '3(&'Hy,,?;<'99yIIu
                    K'3(&';K'3(&H.<!&.&5;(%:('3,/HIIu
                    K'3(&';,&KX3,&HIu
                    K'3(&';9.)*2HIu
                    K'3(&';F.%*&HIu

            {F<(F2HbJF&D(3%,#&Iv
                    &;D'3,(:(<FL1'<F&HIu
            {
        {
{
```

Top-N Precision w/ NN2 Code

```
3ED%'(#Q<6<;3%;+)99&'&5$&<5&'u
3ED%'(#Q<6<;3%;+)99&'&5 h '3(&'u
3ED%'(#Q<6<;3%;V3.&$&<5&'u
3ED%'(#Q<6<;3%;V3.& h '3(&'u

3ED%'(#K&L<;F%'&;M,*(<,F&u
3ED%'(#K&L<;F%'&;M,*(<,F&*u
3ED%'(#K&L<;F.<**393&'*;9),F(3%,*;>).(3.<"&'8&'F&D('%,Z&Ku
3ED%'(#K&L<;F.<**393&'*;b6<.)<(3%,u


D)!.3F#F.<**#1%DZ8'&F3*3%,@#v

        D)!.3F#*(<(3F#6%35#E<3,H:('3,/wx#<'/*Iv

                ('"v
                        M,*(<,F&*#5<(<#j#,&K#M,*(<,F&*H
                                        ,&K#+)99&'&5$&<5&'H
                                                ##,&K#V3.&$&<5&'HyF%E!+V;<'99yIIIu
                        kk*&((3,/#F.<**#<(('3!)(&
                        5<(<;*&(-.<**M,5&JH5<(<;,)E4(('3!)(&*HI#N#?Iu


                        kk#.%<5#),.<!&.&5#5<(<

                        >).(3.<"&'8&'F&D('%,Z&K#E.D#j#,&K#>).(3.<"&'8&'F&D('%,Z&KHIu
                        E.D;*&(0355&,X<"&'*Hy^yIu              #kk#*&(#(2&#2355&,#,%5&*
                        E.D;*&(X&<',3,/$<(&HB;@Iu #kk#*&(#(2&#.&<',3,/#'<(&
                        E.D;*&(1'<3,3,/13E&H[BBIu              #kk#*&(#(2&#('<3,3,/#(3E&
                        E.D;!)3.5-.<**393&'H5<(<Iu###kk#!)3.5#F.<**393&'

                        M,*(<,F&*#),.<!&.&5#j#,&K#M,*(<,F&*H
                                                                        ,&K#+)99&'&5$&<5&'H
                                                                        ##,&K#
V3.&$&<5&'HyF%E!+V;<'99yIIIu
                        kk#*&(#F.<**#<(('3!)(&
                        ),.<!&.&5;*&(-.<**M,5&JH),.<!&.&5;,)E4(('3!)(&*HI#N#?Iu
                        kk#F'&<(&#F%D"
                        M,*(<,F&*#.<!&.&5#j#,&K#M,*(<,F&*H),.<!&.&5Iu
                        kk#.<!&.#3,*(<,F&*
                        M,*(<,F&wx#3,*(<,F&*#j#,&K#M,*(<,F&w?Bxu
                        5%)!.&wx#6<.)&*#j#,&K#5%)!.&w?Bxu

                        9%'H3,(#3#j#Bu#3#z#?Bu#3eeIv
                                3,*(<,F&*w3x#j#,)..u
                                6<.)&*w3x#j#Bu
                        {

                        kk'<,L#3,*(<,F&*#N#%,."#L&&D#(%D#Z
                        9%'#H3,(#3#j#Bu#3#z#),.<!&.&5;,)EM,*(<,F&*HIu#3eeI#v
                        ##5%)!.&#wxF.*X<!&.#j#E.D;53*('3!)(3%,V%'M,*(<,F&H),.<!&.&5;3,*(<,F&H3IIu
                        ##39H6<.)&*wBx#zj#F.*X<!&.w?xI#v#kk.3L&#D'%!<!3.3("
                        ##              9%'H3,(#Q#j#?u#Q#z#?Bu#QeeIv
                                                39H6<.)&*wQx#zj#F.*X<!&.w?x#SS#Q#|j#RIv
```

```
                                        6<.)&*wQN?x#j#6<.)&*wQxu
                                        3,*(<,F&*wQN?x#j#3,*(<,F&*wQxu
                                    {
                                    &.*&#39#HQ#j j#RIv
                                        6<.)&*wQN?x#j#6<.)&*wQxu
                                        3,*(<,F&*wQN?x#j#3,*(<,F&*wQxu
                                        6<.)&*wQx#j#F.*X<!&.w?xu
                                        3,*(<,F&*wQx#j#),.<!&.&5;3,*(<,F&H3Iu
                                    {
                                    &.*&v
                                        6<.)&*wQN?x#j#F.*X<!&.w?xu
                                        3,*(<,F&*wQN?x#j#),.<!&.&5;3,*(<,F&H3Iu
                                        !'&<Lu
                                    {
                                {
                            ##{
                            {

                            .<!&.&5;5&.&(&HIu
                            9%'H3,(#3#j#Bu#3#z#?Bu#3eeI
                                .<!&.&5;<55H3,*(<,F&*w3xIu

                            kk&6<.)<(&#(2&#3,*(<,F&*
                            b6<.)<(3%,#&6<.#j#,&K#b6<.)<(3%,H.<!&.&5Iu
                            &6<.;F'%**T<.35<(&>%5&.H
                            ######E.D#.<!&.&5A#?BA#.<!&.&5;/&($<,5%EZ)E!&'c&,&'<(%'H?IIu
                            :"*(&E;%)(;D'3,(.,H&6<.;(%:)EE<'":('3,/H(')&IIu

                            kkK'3(&#3,*(<,F&*#H(%#5%)!.&#F2&FL#,%#5%)!.&*I
                            +)99&'&5 h '3(&'#K'3(&'#j#,&K#+)99&'&5 h '3(&'H
                            ####################,&K#V3.& h '3(&'Hy,,@;<'99yIIu
                            K'3(&';K'3(&H.<!&.&5;(%:('3,/HIIu
                            K'3(&';,&KX3,&HIu
                            K'3(&';9.)*2HIu
                            K'3(&';F.%*&HIu


                        {F<(F2HbJF&D(3%,#&Iv
                            &;D'3,(:(<FL1'<F&HIu
                        {
                {
{
```

Top-N Precision w/ BN Code

```
3ED%'(#0<6<;3%;+)99&'&5$&<5&'u
3ED%'(#0<6<;3%;+)99&'&5 h '3(&'u
3ED%'(#0<6<;3%;V3.&$&<5&'u
3ED%'(#0<6<;3%;V3.& h '3(&'u

3ED%'(#K&L<;F%'&;M,*(<,F&u
3ED%'(#K&L<;F%'&;M,*(<,F&*u
3ED%'(#K&L<;F.<**393&'*;!<"&*;+<"&*Z&(u
3ED%'(#K&L<;F.<**393&'*;!<"&*;,&(;*&<'F2;/.%!<.;mu
3ED%'(#K&L<;F.<**393&'*;b6<.)<(3%,u


D)!.3F#F.<**#1%DZ8'&F3*3%,^#v

        D)!.3F#*(<(3F#6%35#E<3,H:('3,/wx#<'/*Iv

                ('"v
                        M,*(<,F&*#5<(<#j#,&K#M,*(<,F&*H
                                                ,&K#+)99&'&5$&<5&'H
                                                        ##,&K#V3.&$&<5&'HyF%E!+V;<'99yIIIu
                        kk*&((3,/#F.<**#<(('3!)(&
                        5<(<;*&(-.<**M,5&JH5<(<;,)E4(('3!)(&*HI#N#?Iu


                        kk#.%<5#),.<!&.&5#5<(<

                        +<"&*Z&(#E.D#j#,&K#+<"&*Z&(HIu
                        _@#L@#j#,&K#_@HIu
                        L@;*&(><JZ'P98<'&,(*H?Iu kk#*&(#E<J#}#%9#D<'&,(*
                        E.D;*&(:&<'F24./%'3(2EHL@Iu#kk#*&(#*&<'F2#<./%'3(2E

                        E.D;!)3.5-.<**393&'H5<(<Iu###kk#!)3.5#F.<**393&'

                        M,*(<,F&*#),.<!&.&5#j#,&K#M,*(<,F&*H
                                                                                ,&K#+)99&'&5$&<5&'H
                                                                                ##,&K#
V3.&$&<5&'HyF%E!+V;<'99yIIIu
                        kk#*&(#F.<**#<(('3!)(&
                        ),.<!&.&5;*&(-.<**M,5&JH),.<!&.&5;,)E4(('3!)(&*HI#N#?Iu
                        kk#F'&<(&#F%D"
                        M,*(<,F&*#.<!&.&5#j#,&K#M,*(<,F&*H),.<!&.&5Iu
                        kk#.<!&.#3,*(<,F&*
                        M,*(<,F&wx#3,*(<,F&*#j#,&K#M,*(<,F&w?Bxu
                        5%)!.&wx#6<.)&*#j#,&K#5%)!.&w?Bxu

                        9%'H3,(#3#j#Bu#3#z#?Bu#3eeIv
                                3,*(<,F&*w3x#j#,)..u
                                6<.)&*w3x#j#Bu
                        {

                        kk'<,L#3,*(<,F&*#N#%,."#L&&D#(%D#Z
                        9%'#H3,(#3#j#Bu#3#z#),.<!&.&5;,)EM,*(<,F&*HIu#3eeI#v
                        ##5%)!.&#wxF.*X<!&.#j#E.D;53*('3!)(3%,V%'M,*(<,F&H),.<!&.&5;3,*(<,F&H3IIu
```

```
                    ##39H6<.)&*wBx#z j #F.*X<!&.w?xI#v#kk.3L&#D'%!<!3.3("
                    ##          9%'H3,(#Q# j #?u#Q#z#?Bu#QeeIv
                                39H6<.)&*wQx#z j #F.*X<!&.w?x#SS#Q#| j #RIv
                                        6<.)&*wQN?x# j 6<.)&*wQxu
                                        3,*(<,F&*wQN?x# j #3,*(<,F&*wQxu
                                {
                                &.*&#39#HQ# j j #RIv
                                        6<.)&*wQN?x# j #6<.)&*wQxu
                                        3,*(<,F&*wQN?x# j #3,*(<,F&*wQxu
                                        6<.)&*wQx# j #F.*X<!&.w?xu
                                        3,*(<,F&*wQx# j #),.<!&.&5;3,*(<,F&H3Iu
                                {
                                &.*&v
                                        6<.)&*wQN?x# j #F.*X<!&.w?xu
                                        3,*(<,F&*wQN?x# j #),.<!&.&5;3,*(<,F&H3Iu
                                        !'&<Lu
                                {
                        {
                    ##{
                    {

                    .<!&.&5;5&.&(&HIu
                    9%'H3,(#3# j #Bu#3#z#?Bu#3eeI
                            .<!&.&5;<55H3,*(<,F&*w3xIu

                    kk&6<.)<(&#(2&#3,*(<,F&*
                    b6<.)<(3%,#&6<.# j #,&K#b6<.)<(3%,H.<!&.&5Iu
                    &6<.;F'%**T<.35<(&>%5&.H
                    ######E.D&#.<!&.&5&#?B&#.<!&.&5;/&($<,5%EZ)E!&'c&,&'<(%'H?IIu
                    :"*(&E;%)(;D'3,(..,H&6<.;(%:)EE<'":('3,/H(')&IIu

                    kkK'3(&#3,*(<,F&*#H(%#5%)!.&#F2&FL#,%#5%)!.&*I
                    +)99&'&5 h '3(&'#K'3(&'# j #,&K#+)99&'&5 h '3(&'H
                    ####################,&K#V3.& h '3(&'Hy!,;<'99yIIu
                    K'3(&';K'3(&H.<!&.&5;(%:('3,/HIIu
                    K'3(&';,&KX3,&HIu
                    K'3(&';9.)*2HIu
                    K'3(&';F.%*&HIu

                {F<(F2HbJF&D(3%,#&Iv
                        &;D'3,(:(<FL1'<F&HIu
                {
        {
{
```