

AQLGDVQQIFNQLSDDIATRVAEINALKAQGDAVWPVLSYADIKAGHVTAEQREQIKRRGCAVIKGH
FPREQALGWDQSMILDYLRNRDFDEVYKGGDNFFGTLASRPEIYPIYWSQAQMQRQSEEMANA
QSFLNRLWTFESDGKQWFNPDVSVIYDRIRRRPPGTTSKGLGAHTDSGALERWLLPAYQRFANVF
NGNLAQYDPWHAHRTEVEEYTVDNNTTKCSVFRFTFQGTALSDMLPGQGLLHVPIPEAMAYVLL
RPLDDVPEDELGCVAPGRVLPVSEQWHPLLIEALTSIPKLEAGDSVWWHCDVIHVSAPVENQQGW
GNVMYIPAAPMCEKNLAYAHKVKAALEKGASPGDFPREDYETNWEGRFTLADLNIHGKRALGMDV
1111111111111111111111111111111111110000000000000000000000000000
00
000000001111111111111000
001110000000
00
00
0000000001

>DM_train277

MVKASGIPIALGVWGLCWSLATVNSVPLTSAHGNVTEGESGKPEADVIEQCSDGWSFDATTLDDN
GTMLFFKDEFVWVKSHRGIRELISERWKNFIGVPDAAFRHGHTSVYLIKGDKVWVYTSEKNEKVYPSL
QDEFFGPIPLDAAVECHRGECQDEGILFFQG NRKFWDLTTGTTKERSWPAVGNCT SALRWLGRY
YCFQGNQFLRFPVSGEVPYPGYP LDVRDYFLSCPGRGRSSHRNSTQHGHSTRCDPLVLSAMVS
DNHGATYVFSGSYHWRDLTNRDGHWSWPIAHQWPQPSTVDAAFSWEDKLYLIQDTKVYVFLTK
GGYTLVNGYPKRLEKELGSPVISLEAVDAAFVCPGSSRLHIMAGRRLWWLDLKSGAQATWTELPWP
HEKVDGALCMEKPLGNPNSCSTSGPNLYLIHGPNLYCYRHVDKLNAAKNLPQPQRVSRLLGCTH
11111111111111111111111111111111111110000000000000000000000000
00011111100000000000000000
00
000000000000000111111000
00
00
0000000000

>DM_train278

MKAANASSAEAYRVLSRAFRFDNEDQLLWWHSTAPMFAMLETANYTTPCQYQYLITYKECVIPSL
GCYPTNSAPRWLSILTRYGTPFELS LNCSNSIVRYTFEPINQHTGTDKDPFNTHAIWESLQHLLPLEKSI
DLEWFRHFKHDLTLNSEESAFLAHNDRLVGGTIRTQNKALDLKDGRFALKTYIYPALKAVVTGKTIHE
LVFGSVRRLAVREPRILPPLNMLEEYIRSRGSKSTASPRLVSCDLTSPAKSRIKIYLLQMV SLEAMEDLW
TLGGRRRDASTLEGLSLVRELWDLIQ LSPGLKSYAPYLP LGVIPDERLPLMANFTLHQNDPVPEPQVY
FTTFGMNDMAVADALTTFFERRGWSEMARTYETTLKSYPHADHDKLNLYLHAYISFSYRDRTPYLSVY
LQSFETGDWAVANLSESKVKCQDAACOPTALPPDLSKTGVYYSGLH
1100
00
00
00
00
000111111111111111111
1111111111

>DM_train281

>DM_train475

CEGPPPGTEQIGYRGMENYYVKRQRALSIQANQPVESLPAADSTGPKASEVYQSVQVLKDLVSGE
FTRTMVAVTTWVSPKEGCNYCHVPGNWASDDIYTKVVSRRMFELVRAANSWKAHVAETGVTCYT
CHRGNPVPKYAWVTDPGPKYPSGLKPTGQNYGSKTVAYASLPFDPLTPFLDQANEIRITGNAALAGS
NPASLKQAEWTFGLMMNISDSLGVGCTSCHNTRAFNDWTQSTPKRRTAWYAIRHVRDINQNYIWP
LNDVLPASRKGPYGDPLRVSCMTCHQAVNKPLYGAQMAKDYPGLYKTAVTQEALPGSAPASEAAP
AAATEAAPEAPAQEVPAAEAVPAAAEPGAEEAAGSVEPAPVEEVAPAPAAQRL
00
00
00
00
000000000011
11111111

>DM_train479

MYALTQGRIFTGHEFLDDHAVVIADGLIKSVCPVAELPPEIEQRSLNGAILSPGFIDVQLNGCGGVQFN
DTAEAVSVETLEIMQKANEKSGCTNYLPTLITTSDELKQGVVRVMREYLAKHPNQALGLHLEGPWLN
LVKKGTHNPNFVRKPDAAALVDFLCENADVITKVTLAPEMVPAEVISKLANAGIVVSAGHSNATLKEAK
AGFRAGITFATHLYNAMPYITGREPLAGAILDEADIYCGIADGLHVDYANIRNAKRLKGDKLCVTD
ATAPAGANIEQFIFAGKTIYYRNLGCVDENGTSGSSLTMIEGVRNLVEHCGIALDEVLRMATLYPARA
IGVEKRLGTLAAGKVANLTAFTPDFKITKTIVNGNEVVTQ
1000
00
1100
00
11111100
0000000

>DM_train481

MKWSDNKDDILKGGDVVRLFHAEQEKFLTCDEHRKKQHVFRLRTTGRQSATSATSSKALWEVEVQH
DPCRGGAGYWNSLFRFKHLATGHYLAAEVDPDFEECLEFQPSVDPDQDASRSRLRNAQEKMVVSL
VSVPEGNDISSIFELDPTTLRGGDSLVRNRSYVRLRHLCTNTWVHSTNIPIDKEEEKPVMKIGTSPLKE
DKEAFAIVPSPAVERDLDFANDASKVLGSIAGKLEKGTITQNERRSVTKLLEDLVYFVTGGTNSGQDV
LEVVFSPKPNRERQKLMREQNILKQIFKLLQAPFTDCGDGPMLRLEELGDQRHAPFRHICRLCYRVLRH
SQQDYRKNQEYIAKQFGFMQKQIGYDVLAEADTITALLHNNRK
11111111111000
111000
11111111111100
00
000001111111111111111000
000000

>DM_train485

SLNMFWFLPTHGDGHYLGTEEGSRPVDHGYLQQAADRLGYTGVLIPTRSCEDAWLVAASMIP
VTQRLKFLVALRPSVTSPTVAARQAATLDRLSNGRALFNLTGSDPQELAGDGVFLDHSEYEAFAEF
TQVWRLLQRETVDNFKHIVRGAKLLFPAIQPYPPLYFGSSDVAQELAAEQVDLYLTWGEPPE
LVKEKIEQVRAKAAAHGRKIRFGIRLHVIVRETND EAWQAERLISHLDDETIAKAQAARTDSVQG

KVNVEVDFPGKSTVVLTRPSVFKPLEGAGSPVSNVDAPSVQRSQNKDYVEVGGNDIDITVDFI
MSIGRGIGEETNVEQFRELADEAGATLCCSRPIADAGWLPKSRQVGQSGKVVGSCKLYVAMGISGSI
QHMAGMKHVPTIIAVNTDPGASIFTIAKYGIVADIFDIEELKAQLAA
00
00
00
11
11
11

>DM_train732
KYVPQRKPQTSRRHTPASSSSSSSSSSSSSSSSSSSDSDMTVSAESFEKHSKPKVIVLRAVRADGKQQ
GLQTTLYYGLTSNGLPKAKIVAVELSDLSVWKLCAKFRLSAHMKAKAAIGWGKNCQYRAMLEASTG
NLQSHPAARVDIKWGRLPSSLQRAKNALLEN GAPVIASKLEMEIMPKANQKHQVSVILAAMTPRRM
NIIVKLPKVTYFQQGILLPFTFSPRPFWRPEGSQSDSLPAQIASAFSGIVQDPVASACELNEQSLTTFN
GAFFNYDMPESCYPHVLAQECSSRPPFIVLIKLDERRISLELQLDDK
1100000000000000000000
00
0001
11
11

>DM_train740
MSEEGPVKIREASKDNVDFILSNVDLAMANSRRLRVMIAEIPTLAIDSVEVETNTTVLADDEFIAHRLGLI
PLQSMIEQLEYSRDCFCEDHCDKCSVVLTLQAFGESESTTVYVSKDLVIVSNLMGRNIGHPIIQDKEG
NGVLICKLRKQELKLTCAVAKKIAKEHAKWGPAAIEFEYDPWNKHKHTDYWYEQDSAKEWPQSK
NCEYEDPPNEGDPFDYKAQADTFYMNVESVGSIPVDQVVVRGIDTLQKKVASILLALTQMDQDKVN
FASGDNNTASNMLGSNEDVMMTGAEQDPYNSASQMGNTGSGGYDNAW
00
00
00
00
00
11
11

>DM_train748
MKISIGLGKEGVEERLAERGVSRDFLKFTAIAVMTGMGPAFAPEVARALMGP RRPSVVYLHNAECT
GCSESVLRAFEPYIDTLILDITLSLDYHETIMAAAGDAAEALEQAVNSPHGFI AVVEGGIPTAANGIYGK
VANHTMLDICSRLPKAQAVIAYGTCAATFGGVQA AKPNPTGAKGVNDALKHLGVKAINIAGCPPNPY
NLVGTIVYYLKNKAPELDSLNRPTMFFGQTVHEQCPRLPHFDAGEFAPSESEEARKGWCLYELGCK
GPVTMNNCPKIKFNQTNWPVDAGHPCIGCSEPDFWDAMTPFYQN
11100000000000000000000
00
00
00
00
00

>DM_train765
LSYPNATESVYGLTRYSTNDEAIAGVNNESSITPAKFTVALNNVFETR VSTESSNGVIKISSLPQALAGA
DDTTAMTPLKTQQLAVKLI AQIAPSKNAATESEQVIQLATVAQARQGLTREGY AISPYTFMNSTATE

00000011111111111111111000000000000000000000000000011111111111111111

>DM_train889

DAETVGRVKRFIQENVFVAANHNGSPLSIKEAPKELSFGARAEHPRIHPVASKLLRLMQKKETNLCLSA
DVSLARELLQLADALGPSICMLKTHVDILNDFTLDVMKELITLAKCHEFLIFEDRKFFADIGNTVKKQYEG
GIFKIASWADLVNAHVVPGGVVKGLQEVGLPLHRGCLLIAEMSSTGSLATGDYTRAAVRMAEEHSE
FVVGFIGSRVSMKPEFLHLPQVLEAGGDNLGQQYNPQEVIGKRGSDIIVGRGIISAADRLEAAE
MYRKAWEAYLSRLGV

1111111111111111111111111111111000000000000000000000000000000000000000
00
00
0001

>DM_train894

KYALADASLKMADPNRFRGKDLPLDQLTDPGGVRRVYHIQAGLPDPFPSPITVYAVLERACRS
VLLNAPSEAPQIVRGASEDVRKQPYNLTIAWFRMGNGAIPITVMEYTECSYNKSLGACPIRTQPRW
NYYDSFSAVSEDNLGFLMHAPAFETAGTYLRLVKINDWTEITQFILEHRAKGSCKYALPLRIPPSACLSP
QAYQQGVTVDSIGMLPRFIPENQRTVAVYSLKIAGWHGPKAPYTSTLLPPELSETPNATQPELAPEDP
EDSALLEDPVGTTHHHHH

00
00
00
000

>DM_train901

GAMGMKVDVGPDPPLVYRPDVPDPEMAKSKDSFRNYTSGPLLDREVFTTYKLMHHTHQTVDFVSRKRIQ
YGSFSYKMTIMEAVGMLDDLDVDESDDVPDFPNFHFQTAEGIRKHPDKDFHVLVGLLHDLGKI
MALWGEQWAVVGDTPVGCRCRQASVVFCDDSTFDNDPDLQDPRYSELGMYQPHCGLENVLMMS
WGHDEYLYQMMKFNKFSLPSEAFYMIRFHSPWHTGGDYRQLCSQQDLDMPLPWWQEFNKFDLY
TKCPDLPDVESLRPYQGLIDKYCPGTLSW

1111111111111111111111111111111000000000000000000000000000000000000000
00
00
00

>DM_train902

HHHHHEQKLISEEDLEGIKYRRNELFLDVLESVLLMSPQGQVLSAHVSGRVVMKSYLSGMPECKF
GMNDKIVIEKQKGTADETSIGKQSIAIDDCTFHQCVRLSKFDSERSISFIPDGEFELMRYRTTKDII
PFRVIPLVREVGRTKLEVKVVIKSNFKPSLLAQKIEVRIPTPLNTSGVQVICMKGKAKYKASENAIWKIK
RMAGMKESQISAEIELPTNDKKKWARPPISMNFEVFPFAPSGLKVRYLKVFEPKLNYSHDHVIKWVRYI
GRSGIYETRC

1111111111111111111111111111111000000000000000000000000000000000000000
11
00
000111111111000

>DM_train911

EDRYKEKLLQKAKAEGVESIEELKKRLADQIEEKKELNKIDPLRELEQHNLNAGSRIHTNKEHKTTKMSN
KSNEKSGNVLPKDKPYKTLDDYKLDKIKDLSKQEVFLWRKWSNRDDSLVAVVPYVKTFQGMYYKY

MSRKMSIQEYELIHQDKEDEGLRKYRRQCMQDMHQKLSFGPRYGFVYELETGEQFLETIEKEQKVT
IVVNIYEDGVRGCDALNSSLECLAAEYPMVKFKIRASNTGAGDRFSSDVLPTLLVYKGGELISNFISVA
EQFAEDFFAADVESFLNEYGLLPEREIHDLGQNTTEDEDIE
1111111111110000000000000000000000000111111111111111111111111111111
11111111111100
00
00000111111111111111

>DM_train1152

GALPSHNKASFTDEEFILDVVRKNPTRRTTHLYDEISHYVPNHTGNSIRHRFRVYLSKRLEYVYEVD
KFGKLVRRDDGNLIKTKVLPSPSIKRFSADEYTLAIAVKKQFYRDLFQIDPDTGRSLTDEDPTAIARR
NMTMDPNHVPGSEPNFAAYRTQSRRGPIAREFFKHFAEEHAAHTENAWRDRFRKFLAYGIDDYISY
YEA EKAQNREPEPMKNLTNRPKRPGVPTPGNYNSAAKR
11111100
000111111111111111111
1111111111000111111100000
01111111100000001111

>DM_train1177

GSHMQGQSPIKQEMKELIRRSQEGDQEARDEIIEKNMRLVWSVQRFLNRGYEADDLFQIGCIGLL
KSVDKFDLSYDVKFSTYAVPMIIEIQRFLRDDGTVKVSRSLKEMGNKIRKAKDEL SKTRGRAPTVTEIA
DHLGISPEDVVLAEAVRSLPHTVYENDGDPITLLDQIADAEASWFDKIALKKAIEELDERERLIVY
LRYKQDTQSEVASRGLISQVQMSRLEKILQH IK
11
111111111111111111111100
00000011
11111111111111111

>DM_train1184

HVKLEQLGIPRQVLYKDDFLKSTQKSPATAPKPMPPRPPLPNQQFGVSLQLHLEKNPEQEPPIVLR
VAYLQAHALTTTEGIFRRSANTQVVREVQKYNMGLPVDFDQYNELHLPVILKTFRELPEL LTFDLY
PHVVGFLNIDESQRPATLQVLQTLPEENYQVLRFLTAFLVQISAHSDQNKMTNTNLAVVFGPNLLW
AKDAAITLKA INPINTFTKFLLDHQGELFPSPDPSGL
1111111111111111111111111111111111000000000000000000000000000000000000
00
00011111111100000000
0000000011111111

>DM_train1197

ADNEEDVEAEEY TELTDISGVGPSKAESLREAGFESVEDVRGADQSALADVSGIGNALAARIKADVG
GLEVESETEAEVEEGGEEAPDEDVETELQARGLTEKTPDLSE DARLLTQRHRVGKPKQFNRQDHHKK
KRVSTSWRKPRGQLSKQRRGIKKGDTVEAGFRSPTAVRGKHPSGFEEVRVHNVDDELEGVDGDTEA
VRIASKVGARKRERIEEAAEDAGIRVLNPTYVEVEVSE
11
11111111111111111100
00
00000000001111

>DM_train1203

1

>DM_train2028

MYHHYHAFQGRKLTQERARVLEFQDSIHYSRYSDDNYEYRHVMLPKAMKLVIPSDYFNSEVGLR
ILTEDEWRGLGITQSLGWEHYECHAPEPHILLFKRPLNYEAELEAATAAAQEQEQEQEQEQEQEQEQEQ
QQHQQTQSI SNMQVPPQIS

11111000
000000000000000000000000000000000000011111111111111111111111111111111

>DM_train2030

SDNDDIEVESDADKRAHHNALERKRRDHIKDSFHSRLRDSVPSLQGEKASRAQILDKATEYIQYMRRKN
HTHQQDIDDLKRQNALLEQQVRALEKARSSAQLQTNYPSSDNSLYTNAKGSTISAFDGGSDSSSESE
PEEPQSRKLRMEAS

1111111100
000000111

>DM_train2049

MDLEKNYPTPTSRTGHGGVNQLGGVFNGRPLPDVVRQRIVELAHQGVVRPCDISRQLRVSHGCVS
KILGRYYETGSIKPGVIGGSKPKVATPKVVEKIAEYKRQNPTMFAWEIRDLLAERVCDNDTVPSVSSIN
RIIRTKVQQPPNQ

11
11111110011111111

>DM_train2084

MFGRDPFDLRFERMFKEFFATPMTGTTMIQSSTGIQISGKGFMPISIIEGDQHIKVIAWLPGVNKEDIIL
NAVGDLEIRAKRSPLMITESERIIYSEIPEEEEEYRTIKLPATVKEENASAKFENGVLVILPKAESSIKKGINI
E

11100000000000000000000000000000
00

>DM_train2094

MQVILLEPSRLGKTGEVVSVKGDGYARNWLPQGLAVSATRTNMKTLEAQLRSIEKRQAQEKAVAEDL
ASRLNGVAVELSVRAGEGKIYGAVTHQDVANSLDQLGFDVDRRDKIDMPKTVKEVGEYDIAYRAHPEV
TIPMKLVVHAAK

0011
11

>DM_train2098

KKVVAVVKLQLPAGAATPAPPVGPALGQHGANIMEFVKAFNAATANMGDAIVPEITAYADRSFTFV
TKTPPASYLIRKAAGLEKGAHKPGREKVGRITWEQVLEIAKQKMPDLNTDLEAAARMIAGSARSMGV
EVVGAPEVKDA

00
0011

>DM_train2112

SKEPLRPCR PINATLAVEKEGCPVCITVNTTICAGYCPTMTRVLQGVLPALPQVVCNYRDVRFESIRLP
GCPRGVNPVVSVAVALSCQCALCRRSTTDCGGPKDHPLTCDPRFQDSSSSKAPPSLPSRPLPGP
SDTPILPQ

00
000000000000000000000000000000000000011111111111111111111111111111111

>DM_train2149

GAAKTRKTTTTPRKPRGRPKKLEKEEEEGISQESSEEEQ
111
111

>DM_train2598
MDAQTRRRERRAEKQAQWKAANPLLVGVSAPVNLPLSLNRKPKSRVESALNPIDLTVLAEYHKQIE
SNLQRIERKNQRTWYSKPGERGITCSGRQKIKGKSIPLI
111
111

>DM_train2652
MSSLPFVFGAAASSRVVTAATAAAGKTAETKQEKSFVDWLLGKITKEDQFYETDPIRGGDVKSSGSTSG
KKGTTSGKKGTVSIPSKKNGNGGVFGGLFAKKD
111
111

>DM_train2665
FDLNDFLEQLRQMKNMGGMASLMGKLPGMGQIPDNVKSQMDDKVLVRMEAIINSMTMKERAKP
EIIKGSRRRIAAAGSGMQVDVNRLKQFDDMQRMMKMM
000000000111111111111111111111111111111111000
00000000000000000000000000000000

>DM_train2798
GHMHVWPVQDAKARFSEFLDACITEGPQIVSRRGAEAVLVPIGEWRRRLQAAARPSLKQLLLSDSAR
TEMLVPERGKARRRQVEPLRGS
111001111111111111111111111
1111111111111111

>DM_train2800
DSKYSQTSHKLVQLLTTTAEQQLRHADIDTSCKDVLSCTGTSNSASANS SGGSCPSSHSLTERHKILH
RLQE GSPSDITTL SVEPD
111110000000000000011100000000000000
00000011111111

>DM_train2805
MEKPYKPECGKSFSSNLQKHQRTHHTGEKPYKPECGKSFSSDLQKHQRTHHTGEKPYKPECG
KSFSDHL SRHQRTHQNKK
1110000000000000000000
000000001111

>DM_train2816
MNLEPPKAEIRSATRVMMGGPVTPRKGPPKFQRQRQFQKSPPKKGVQGFQDDIPGMEGLGTDITVI
CPWEAFNHLELHELAYGII
111
111

>DM_train2836
GSESGDNVRSSAGAVRDAGGAFGKREQAEEERYFRARAKEQLAALKKHENEISHHAKEIERLQKEIE
RHKQSIKCLKQSEDDDD
111001111111111111111111111
11111111111

>DM_train2840

