

## **2.9 DATA TO DRIVE ANI Description and example of the b.in file:**

1. PARAMETER CONTROLS THE OUTPUT STREAM, (by selecting numbers from zero to 8, one can include various additional output information, to be printed into output file b.out, also permanently attached under mathematical number NT=2).  
DEBUG=  
**3**
2. NUMBER OF DIFFERENT DATA CLASSES TO BE HANDLED AND NUMBER OF BUTSTRAP REPLICAS (L and NBUT numbers):  
**2, 10**
3. TRAINING SAMPLES NAMES:  
***KASCADE1000***  
***KASCADE1000***
4. CONTROL(EXPERIMENTAL) SAMPLE NAME:  
***KASCADE1000S***
5. DUMP ( ARCHIVE) SAMPLE NAME (name of the file to store current training sample with applied cuts and selections).  
***NN-EST***
6. PAW HBOOK NAME:  
***b.hbook***
7. TOTAL NUMBER OF VARIABLES IN DATA FILES (dimensionality of training and control samples):  
**5**
8. THE RELATIVE COORDINATES (coordinate of the first event)FOR EACH DATA FILE:  
***0,0***
9. TOTAL NUMBER OF EVENTS TO BE READ FROM EACH DATA FILE:  
***500,500***
10. FIRST EVENT COORDINATE and SIZE OF CONTROL (EXPERIMENTAL) DATA FILE:  
***0,100000***
11. STATUS (DENCURVE - producing numerous PAW plots):  
**DENCURVE**
12. OPERATION MODE\*:  
(JMODE = DIMDIM, BHATA, ONE DIMEN., COVCOR, FWRITE, BETEST, ONE-LEAVE-OUT-, CLASSIFICATION, REGRO, BUTSTRAP, LEARNING, CLASSIFI, FAST, SUPERCUT, SOBOL-CUT, MULTI-CUT)  
***LEARNING***
13. THE TYPE OF DATA FILES(ASCII with header and weights, ASCII without HEADER, OR PAW NTUPLES, (ACCESS = SEQUENTIAL, SIMPLE, NOMAD):  
***SEQUENTIAL***

14. DENSITY ESTIMATION MODE. Two general nonparametric modes are implemented, kernel density estimator and K nearest neighbours estimator.

(JDEN = PARZ or KNN):

*PARZ*

15. WEIGHTS IN REGRO MODE ( JDIST = LINEAR, SQUARE, UNIFORM):

*UNIFORM*

16. FORMAT OF SEQUENTIAL INPUT (FORMAT of ASCII string):

*(10F10.5)*

17. NUMBER OF DIFFERENT A PRIORY PROBABILITIES AND LIST OF PROBABILITIES:

(IAP number and AP array):

**3**

**0.5, 0.5**

**0.1, 0.9**

**0.01, 0.99**

18. RECONSTRUCT FIRST TYPE EVENTS PORTION?:

*RECONSTRUCTT*

19. NUMBER AND VALUE OF NUCLEI WIDTHS (OR LIST OF NEAREST NEIGHBORS):

(KCL number and F array)

**5**

**0.3, 0.4, 0.5, 0.6, 0.7 or 15, 25, 50, 100, 150**

20. MAXIMAL EXPONENT IN PARZEN DENSITY ESTIMATION AND STRANGNESS CRITERIUM IN BAYESIAN DECISION RULE: (expmax and Strange):

**9000000., 0.000000000000000000000001**

21. NUMBER OF NEAREST NEIGHBOURS FOR DIMDIM AND REGRO MODES, NUMBER OF PRINCIPAL COMPONENTS IF PCA MODE SELECTED (NEI=):

**17**

22. VARIABLES TO BE PROCESSED (AMOUNT AND RELATIVE NUMBERS) (N number and NUMB array):

**2**

**3, 4**

23. THE MINIMAL DIMENSION OF BEST VARIABLES SUBSET TO BE CHOSEN BY BHATA SUBROUTINE (INTDIM=):

**1**

24. LOWER BOUNDS OF VARIABLES (AMIN array):

**-9999999, -9999999**

25. UPPER BOUND (AMAX array):

**9999999, 9999999**

26. Random generator used (genert = pseudo, or lp-tau - a uniform sieve in N-dimensions):

In BETEST mode also RANNOR, NORRAN and NORMCO generators are used,

***Pseudo***

27. DATA TRANSFORMATION TYPE (normalisation to 0-1 - renorm, principal component transformation -pea), :

**norenorm**

28. PARAMETERS OF PAW HBOOKS (FOR READ AND WRITE):

(ntupw (r) - opening code; ntinw(r) - ntuple or histogram ID, MEMw(r) - memory size, IFOw(r) - hbook ID)

**1,10,100000,11**

**0,20,1000000,22**

29. ID of Ntuple from which input data is downloaded in the NOMAD mode

**1,2**

30. NEURAL NET CONFIGURATIONS OF LAYERS, N OF NODES IN EACH LEYER. (LEYERS number, NODES array):

**3,2,5,1**

31. N OF ITERATIONS, STEP VALUE, SIGMA CRITERIUM,

INITIAL SPREED, RANDOM GENERATOR SHIFT (NITER, cf,sim,spread,ishift numbers):

**3000,100,0.02,88**

32. SPEED (the power index for weighting the difference between actual and desired NN output). WSPEED (the power index for selecting event's weight variants):

$sss(jj)=sss(jj)+abs((o(jj)-b(its+ii)))^{**}speed^{*}u(ITS)^{**}wspeed$

**1.,1.**

33. QUALITY FUNCTION SYMMETRIZATION WEIGHTS (WIGHT array):

**0.5,0.5**

34. SEARCH MODE (search = single - one dimensional search; MULLTI - all net parm. modiflicated simultaneously; neuron - all couplings and threshold of a randomly selected neuron)

**neuron**

35. QUALITY FUNCTION TYPE (qualit = montec - training with M.C.; sigmaa - with ON/OFF pairs; estima - neural estimation:

**montec**

36. QUALITY FUNCTION MODE, qtype = msd, (massa, kolm modes - now suspended)

**msd**

37. MEMORY TYPE (memory = simple - no memorization of better point during search; memory - the best changes are accumulated)

**memory**

38. BEGIN RANDOM SEARCH FROM (begin = random point; or - better point, found in

previous search cycles):

**random**

39. STOP ITERATIONS IF QUALITY FUNCTION IS LESS THAN (stiter number):

**0.00001**

40. DECISION POINT (for 2 class case if qualit ne.montec):

else if(qualit.eq.'sigmaa'.or.qualit.eq.'spectr'.or.qualit.eq.'pure') then  
if(o(1).le.dpoint) then...

**0.51**

41. number of different partitioning of last neuron output (0-1) interval (analog of a priory probabilities for the neural decision making) and list of partitioning and goal functions for all classes.

nipr number and part and goal arrays):

**3**

**0.5, 1**

**0.1,0.9**

**0.1,1.**

42. MULTIDIMENSIONAL "BINS" NUMBER FOR FEATURE SPACE SCANNING

(revealing of multidimensional nonlinear cluster shape), ndel number:

**1000**

43. ESTIMATION MODE: NUMBER OF REGRESSANDS (ONE OR TWO) AND IT'S

NAMES (ny number and numreg array):

**1**

**MUON**