ML 2020 Midterm Problem

1. DS_{new1} =[ ]; use dataset UCI Letters recognition

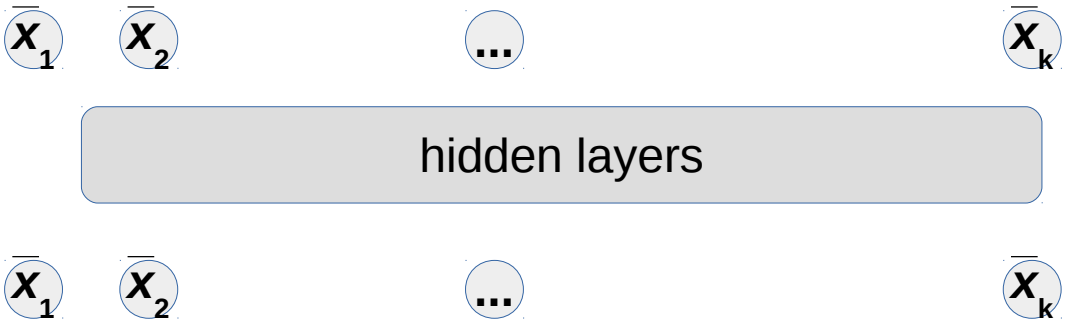
   for all i: 1..k

   for all s in DS : $s_{i=?}$= x_1… x_i = ? … x_k *# set x_i=? , one ? per sample*

   DS_{new1} +=$s_{i=?}$;

2. Split DS (and DS_{new1}) into TS(80%) and TestSet(20%) (TS_{new1}, $TestSet_{new1}$)

3. create *MLP*



4. Train *MLP* with input TS_{new1} and output *TS*
5. Test and evaluate over TestSet (input $TestSet_{new1}$,output TestSet)

ML 2020 Midterm Problem

for all *s* in DS: $s=x_1,\ldots,x_k \rightarrow \underline{x_1},\ldots,\underline{x_k}$ *# encode DS with characteristic vector DS → DS*

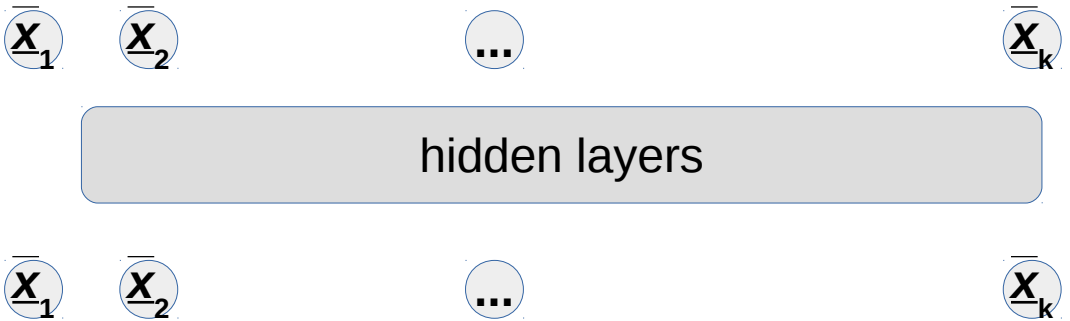
Split DS into TS(80%) and TestSet(20%) TS_{new} =[ ];

for all i: 1..k

for all $s= \underline{x_1}\ldots\underline{x_k}$ *in TS :* $\underline{s_{i=0}} = \underline{x_1}\ldots \underline{x_i} = \underline{0} \ldots \underline{x_k}$ *# set* $\underline{x_i}=\underline{0}$

$\underline{TS}_{new}$ +=$\underline{s_{i=0}}$;

create MLP



Train *MLP* with input $\underline{TS}_{new}$ *and output* TS

1. Test and evaluate over TestSet
2. Test retreiving probabilities...

ML 2020 Midterm Problem

You are not done yet! if you have time try with *k* “?” per sample: DS_{newk}

Also try to train on TS_{new1} and test on $TestSet_{newk}$

*o*r maybe train on TS_{newj} and test on $TestSet_{newk}$

Write a report with all the experimentations and motivate your choices about what to experiment… Due Thu Dec 3.