

Hashing Classification for charged particle tracking

Luiza Adelina Ciucu (ATLAS)

22 May 2020



**UNIVERSITÉ
DE GENÈVE**

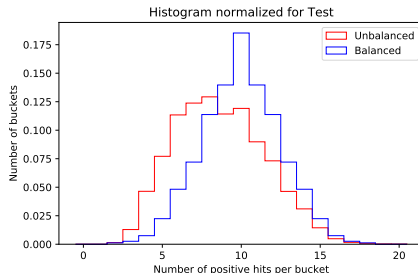
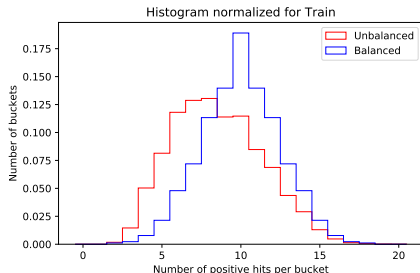
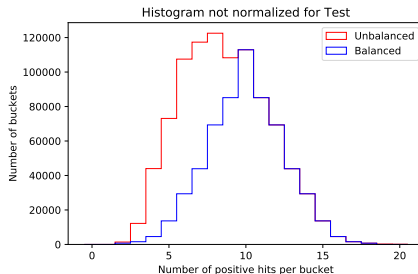
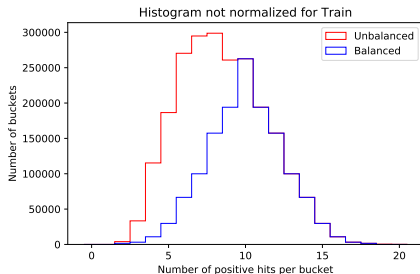
FACULTY OF SCIENCE
Physics Section

Introduction

- Compare two methods to balance the number of positive and negative hits : unbalanced (A) and balanced (B).
- Used 30 events. Each group of 10: 7 in train, 3 in test.
- Unbalanced (A): keep all buckets (with default weight of 1.0).
- Balanced (B): remove buckets so that bucket distribution is symmetric around 10 in NbPositiveHit per bucket (with default weight of 1.0).
- Used min 0, 4, 7 and 10 positive hits in the bucket; if less, consider all hits in the bucket to be negative, but use same reweighting as above.
- How a bucket is created:
 - Loop over events, and for each event build annoy index and then:
 - Loop over hits and for each hit build a bucket using annoy and 20 nearest neighbours by direction
 - Loop over hits in the bucket, find their particleID; find particleID with most hits in the bucket; denote it majority particle.
 - Loop over hits in the bucket again, if belongs to the majority particle assign output +1, else -1.

Histogram NbBuckets vs NbPositiveHit in a bucket

Unbalanced (Balanced) are not (are) symmetric around 10.



Count buckets in each NbPositiveHit category.

- Balanced around 10 NbPositiveHit via bucket removal.
- 0 and 1 no buckets → 19 and 20 are balanced to no buckets (ignored).

```
Original unbalanced. Train:
nbPositiveHit=0 counterBucket=0 percentBucket=0.0
nbPositiveHit=1 counterBucket=0 percentBucket=0.0
nbPositiveHit=2 counterBucket=3973 percentBucket=0.2
nbPositiveHit=3 counterBucket=33349 percentBucket=1.5
nbPositiveHit=4 counterBucket=115266 percentBucket=5.0
nbPositiveHit=5 counterBucket=186516 percentBucket=8.1
nbPositiveHit=6 counterBucket=278380 percentBucket=11.8
nbPositiveHit=7 counterBucket=294847 percentBucket=12.9
nbPositiveHit=8 counterBucket=298765 percentBucket=13.0
nbPositiveHit=9 counterBucket=268813 percentBucket=11.4
nbPositiveHit=10 counterBucket=262604 percentBucket=11.5
nbPositiveHit=11 counterBucket=194065 percentBucket=8.5
nbPositiveHit=12 counterBucket=157420 percentBucket=6.9
nbPositiveHit=13 counterBucket=99938 percentBucket=4.4
nbPositiveHit=14 counterBucket=66595 percentBucket=2.9
nbPositiveHit=15 counterBucket=29752 percentBucket=1.3
nbPositiveHit=16 counterBucket=18828 percentBucket=0.5
nbPositiveHit=17 counterBucket=3184 percentBucket=0.1
nbPositiveHit=18 counterBucket=1560 percentBucket=0.1
nbPositiveHit=19 counterBucket=379 percentBucket=0.0
nbPositiveHit=20 counterBucket=148 percentBucket=0.0
```

UU-:---F1 a.txt All L1 (Text)

```
Original unbalanced. Test:
nbPositiveHit=0 counterBucket=0 percentBucket=0.0
nbPositiveHit=1 counterBucket=0 percentBucket=0.0
nbPositiveHit=2 counterBucket=1383 percentBucket=0.1
nbPositiveHit=3 counterBucket=12281 percentBucket=1.3
nbPositiveHit=4 counterBucket=43997 percentBucket=4.6
nbPositiveHit=5 counterBucket=73892 percentBucket=7.7
nbPositiveHit=6 counterBucket=187498 percentBucket=19.3
nbPositiveHit=7 counterBucket=117346 percentBucket=12.4
nbPositiveHit=8 counterBucket=122522 percentBucket=12.9
nbPositiveHit=9 counterBucket=108282 percentBucket=11.4
nbPositiveHit=10 counterBucket=112926 percentBucket=11.9
nbPositiveHit=11 counterBucket=85146 percentBucket=9.0
nbPositiveHit=12 counterBucket=69333 percentBucket=7.3
nbPositiveHit=13 counterBucket=43918 percentBucket=4.6
nbPositiveHit=14 counterBucket=29392 percentBucket=3.1
nbPositiveHit=15 counterBucket=13665 percentBucket=1.4
nbPositiveHit=16 counterBucket=4578 percentBucket=0.5
nbPositiveHit=17 counterBucket=1578 percentBucket=0.2
nbPositiveHit=18 counterBucket=722 percentBucket=0.1
nbPositiveHit=19 counterBucket=220 percentBucket=0.0
nbPositiveHit=20 counterBucket=122 percentBucket=0.0
```

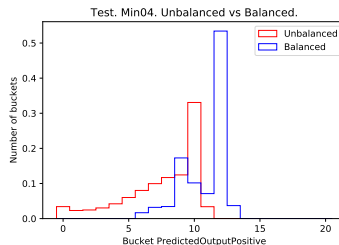
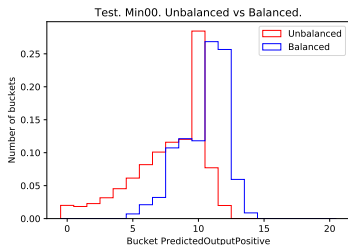
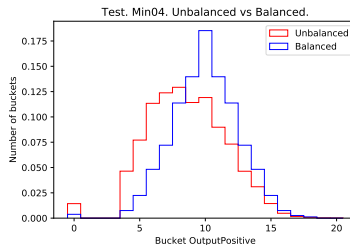
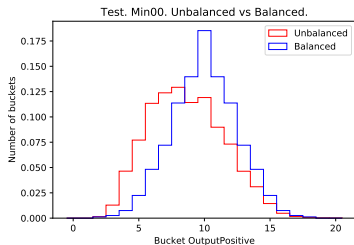
```
BalancedObtained. Train:
nbPositiveHit=0 counterBucket=0 percentBucket=0.0
nbPositiveHit=1 counterBucket=0 percentBucket=0.0
nbPositiveHit=2 counterBucket=1560 percentBucket=0.1
nbPositiveHit=3 counterBucket=3184 percentBucket=0.1
nbPositiveHit=4 counterBucket=18828 percentBucket=0.5
nbPositiveHit=5 counterBucket=29752 percentBucket=1.3
nbPositiveHit=6 counterBucket=66595 percentBucket=2.9
nbPositiveHit=7 counterBucket=99938 percentBucket=4.4
nbPositiveHit=8 counterBucket=157420 percentBucket=6.9
nbPositiveHit=9 counterBucket=194065 percentBucket=8.5
nbPositiveHit=10 counterBucket=262604 percentBucket=11.5
nbPositiveHit=11 counterBucket=194065 percentBucket=8.5
nbPositiveHit=12 counterBucket=157420 percentBucket=6.9
nbPositiveHit=13 counterBucket=99938 percentBucket=4.4
nbPositiveHit=14 counterBucket=66595 percentBucket=2.9
nbPositiveHit=15 counterBucket=29752 percentBucket=1.3
nbPositiveHit=16 counterBucket=18828 percentBucket=0.5
nbPositiveHit=17 counterBucket=3184 percentBucket=0.1
nbPositiveHit=18 counterBucket=1560 percentBucket=0.1
nbPositiveHit=19 counterBucket=0 percentBucket=0.0
nbPositiveHit=20 counterBucket=0 percentBucket=0.0
```

UU-:---F1 b.txt All L1 (Text)

```
BalancedObtained. Test:
nbPositiveHit=0 counterBucket=0 percentBucket=0.0
nbPositiveHit=1 counterBucket=0 percentBucket=0.0
nbPositiveHit=2 counterBucket=722 percentBucket=0.1
nbPositiveHit=3 counterBucket=1578 percentBucket=0.2
nbPositiveHit=4 counterBucket=4578 percentBucket=0.5
nbPositiveHit=5 counterBucket=13665 percentBucket=1.4
nbPositiveHit=6 counterBucket=29392 percentBucket=3.1
nbPositiveHit=7 counterBucket=43918 percentBucket=4.6
nbPositiveHit=8 counterBucket=69333 percentBucket=7.3
nbPositiveHit=9 counterBucket=85146 percentBucket=9.0
nbPositiveHit=10 counterBucket=112926 percentBucket=11.9
nbPositiveHit=11 counterBucket=85146 percentBucket=9.0
nbPositiveHit=12 counterBucket=69333 percentBucket=7.3
nbPositiveHit=13 counterBucket=43918 percentBucket=4.6
nbPositiveHit=14 counterBucket=29392 percentBucket=3.1
nbPositiveHit=15 counterBucket=13665 percentBucket=1.4
nbPositiveHit=16 counterBucket=4578 percentBucket=0.5
nbPositiveHit=17 counterBucket=1578 percentBucket=0.2
nbPositiveHit=18 counterBucket=722 percentBucket=0.1
nbPositiveHit=19 counterBucket=0 percentBucket=0.0
nbPositiveHit=20 counterBucket=0 percentBucket=0.0
```

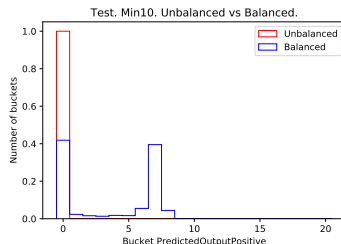
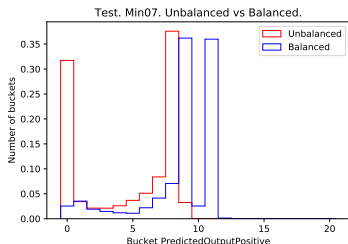
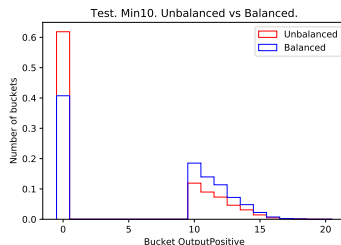
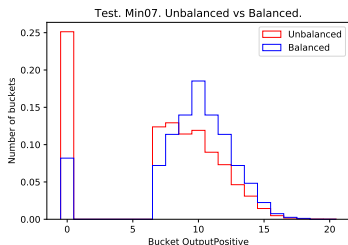
OutputPositive and PredictedOutputPositive 1/2

- Min00 (left) and Min04 (right)



OutputPositive and PredictedOutputPositive 2/2

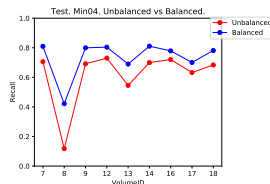
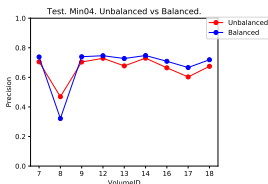
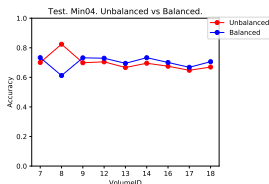
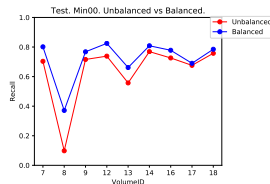
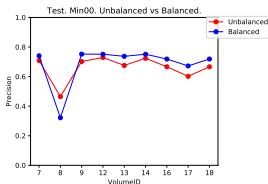
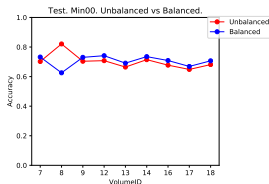
- Min07 (left) and Min10 (right).
- Min10 and method Unbalanced predicts all hits to be negative.



Metrics for each VolumeID overlay two methods 1/2.

- Min00 and Min04 very similar.

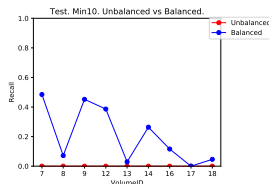
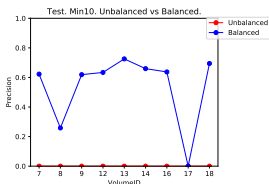
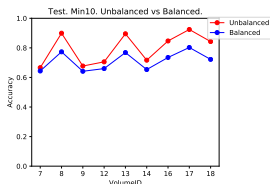
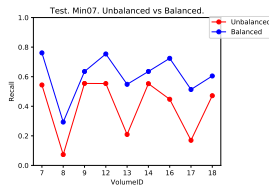
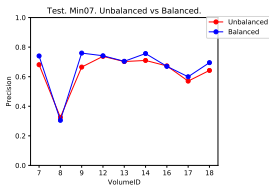
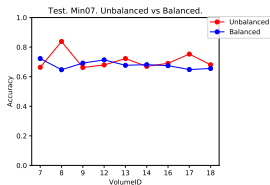
Accuracy	Precision	Recall
$\frac{TP+TN}{TP+FP+FN+TN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$



Metrics for each VolumeID overlay two methods 2/2.

-
- Min10 learns all hits to be negative, so precision and recall at zero.

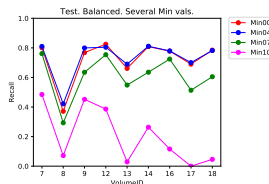
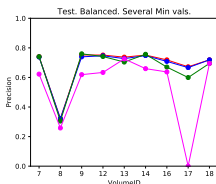
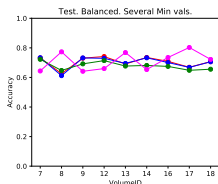
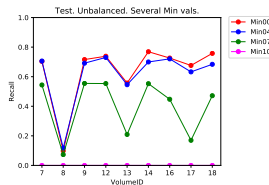
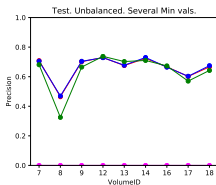
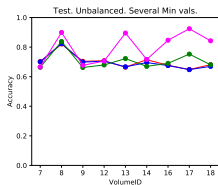
Accuracy	Precision	Recall
$\frac{TP+TN}{TP+FP+FN+TN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$



Metrics for each VolumeID with min value used.

- Min00 and Min04 very similar.
- Min10 learns all hits to be negative, so precision and recall at zero.

Accuracy	Precision	Recall
$\frac{TP+TN}{TP+FP+FN+TN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$



Conclusion. Future plans.

○ Conclusions:

- Compared two methods: Unbalanced vs Balanced.
- Balanced: remove buckets such that number of buckets with a given nbPositiveHit is symmetric around 10.
- 0 and 1 no buckets \rightarrow 19 and 20 are balanced to no buckets (ignored).
- For each Unbalanced and Balanced, did Min00, Min04, Min07, Min10.
- As seen before, Unbalanced Min10 learns all hits to be negative, so precision and recall are zero.
- Min00 and Min04 are very similar.
- Min07 in between Min04 and Min10.
- **Overall, balancing buckets improve performance.**

○ Future plans:

- Balance buckets in majority particle η . Is it replacing the balancing of nbPositiveHit, or in addition to it?