Another view of the same thing: We can have really good predictions, but either because it's a float number (almost never exactly the very same value as the target), it doesn't matter the order (rank) to get a better corr as I have always understood that numerai does:

```
import numpy as np import pandas as pd from numerai_tools.scoring import numerai_corr
```

```
def compare(pred, target): pred = pd.DataFrame({'predictions': pred}) target = pd.DataFrame({'target': target})['target'] return numerai_corr(pred, target)
```

```
target = [0] + [.5]*100 + [1] pred = [0] + list(np.linspace(0.45,0.55,len(target)-2)) + [1] out = compare(pred, target) print(f'output: {out[0]:.3f}')
```

```
for _ in range(10): pred = [0] + list(np.random.uniform(low=0.45, high=0.55, size=100)) + [1] out = compare(pred, target) print(f'output: {out[0]:.3f}')
```

Output

output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468 output: 0.468

So, the order around that 0.5 value of the target, does not make a difference (here a random cloud of points near the target value was generated each time).

I must be misunderstanding something, I suspect:

1. That the actual score is done against a non-discretized version of the target, which would mean that we have an artificially difficult target (bad for everyone, isn't it?)

2. Maybe I'm simplifying the example too much? But the target is 5 buckets… I'm working around one. I think even if the "output" changes in value, the effect is the same.

If it is none of those, I wonder if this message still applies:

[Are predictions discrete or continuous?](#)

Nevertheless, your predictions should be real valued and ideally not contain any ties (they will be broken by row order, i.e. essentially randomly).

Because according to my analysis, breaking the tie drops the corr from that 0.9ish to 0.5ish values (again in this example).

Any idea of what I'm missing?