Another curiosity, how do top staked models (models != accounts) perform ?

[

cumulative

1200×800 80.1 KB

](https://forum.numer.ai/uploads/default/original/2X/8/844b251c7da7ffeabc9052a1a0c16cd846f114cf.png)

Download the data, then plot

!/usr/bin/env python3

import sys import pandas as pd import numpy as np import matplotlib.pyplot as plt import seaborn as sns

if len(sys.argv) < 2: print("Usage:") print(f" {sys.argv[0]} round-xxx-yyy.csv") sys.exit(1)

CORR_COL='v2Corr20'

df = pd.read_csv(sys.argv[1])

keep only the needed columns

df = df[["corrWMetaModel","mmc","modelName","payoutSettled","selectedStakeValue","tc","v2Corr20","roundNumber"]]

keep only staked models

df = df[df["selectedStakeValue"] > 1.0]

make sure of the sorting

df = df.sort_values(by="roundNumber", ascending=True)

plt.rcParams["figure.figsize"] = [12,8] # default is [6.4, 4.8]

def plot(s): ax = ((s.fillna(0.) + 1.0).cumprod() - 1.0).plot(kind='line', legend=True, linewidth=3) return ax

tmp_series = df.groupby(['roundNumber']).apply(lambda x: x[CORR_COL].mean()) tmp_series.name='All staked models Mean v2Corr20' plot(tmp_series)

tmp_series = df.groupby(['roundNumber']).apply(lambda x: (x[CORR_COL]*x["selectedStakeValue"]).sum()/x["selectedStakeValue"].sum()) tmp_series.name='Stake Weighted v2Corr20' plot(tmp_series)

tmp_series = df.groupby(['roundNumber']).apply(lambda x: x.nlargest(3, "selectedStakeValue")[CORR_COL].mean()) tmp_series.name='Highest 3 Staked Model Mean v2Corr20' plot(tmp_series)

tmp_series = df.groupby(['roundNumber']).apply(lambda x: x.nlargest(10, "selectedStakeValue")[CORR_COL].mean()) tmp_series.name='Highest 10 Staked Model Mean v2Corr20' plot(tmp_series)

tmp_series = df.groupby(['roundNumber']).apply(lambda x: x.nlargest(30, "selectedStakeValue")[CORR_COL].mean()) tmp_series.name='Highest 30 Staked Model Mean v2Corr20' ax = plot(tmp_series)

ax.get_figure().savefig(f"cumulative.png")