**Varying n\_estimators (20 through 5000)**

Graphical user interface

Description automatically generated

**Varying n\_estimators (1 through 20)**

Best metrics are at 16, BUT this means that all predictions are around 15 (plus or minus 1 or 2) – convenient middle grounds = predictions are not diverse enough!

500 seems to be too much (overfitting?) – there is diversity in predictions, but none of them match the ground truth

200 – some close calls, but not as many actual matches as visually confirmed for 100

100 – best n\_estimators?

1000 – also a lot of actual matches, on the other hand

Graphical user interface

Description automatically generated with medium confidence

**Varying LR (0.01 through 1 with a step 0.01)**

Calendar

Description automatically generated

**Varying gamma (0.1 through 10 with a step 0.1)**

Text, calendar

Description automatically generated

**Varying gamma (10 through 1000 with a step 5)**

Timeline

Description automatically generated with low confidence

**Varying gamma (1000 through 10000 with a step 50)**

First positive R2 at gamma=1250

Calendar

Description automatically generated

**Reg\_alpha (0.1 through 1, step 0.1)**

Disregard gamma on plot

Website

Description automatically generated

**Reg\_lambda (0.1 through 1, step 0.1)**

Chart, line chart

Description automatically generated with medium confidence

**Reg\_lambda (1 through 10, step 1)**

Website

Description automatically generated

**N\_steps**

Calendar

Description automatically generated

**Subsample**

Graphical user interface

Description automatically generated

**colsample\_bylevel**

A picture containing text

Description automatically generated

**colsample\_bynode**

Calendar

Description automatically generated

**colsample\_bytree**

A picture containing text

Description automatically generated