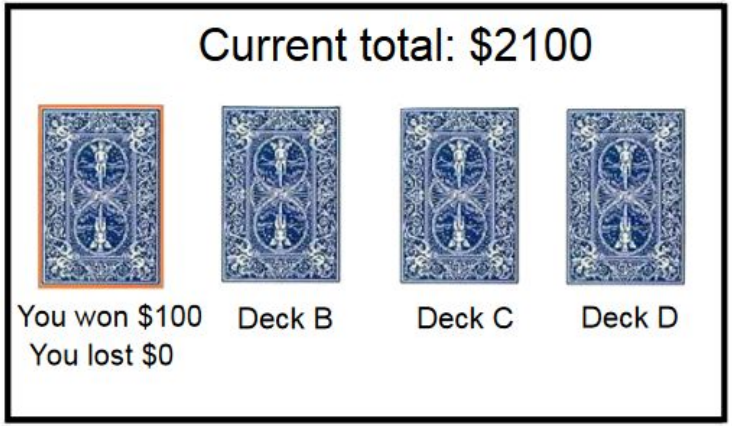
OED - Iowa Gambling Task

Anish Haluvani Sundresh, Jennifer Kempkens,
Hannah Köster, Rowan Mohamed, Jessica Simon

Introduction - Experimental Framework

- Data from 617 healthy participants performing the Iowa Gambling Task
- Characteristics
  - a psychological task thought to simulate real-life decision making
  - Variables
    - current reward
    - choice
    - differences in papers,
    - number of trials,
    - payoff scheme



Introduction - Experimental Framework

- Existing dataset: Iowa Gambling Task

Study	Number of participants	Number of trials	Payoff	Demographics[a]
Fridberg et al. [3]	15	95	1	M = 29.6 years (SD = 7.6)
Horstmann[b]	162	100	2	M = 25.6 years (SD = 4.9), 82 female
Kjome et al. [5]	19	100	3	M = 33.9 years (SD = 11.2), 6 female
Maia & McClelland [6]	40	100	1	Undergraduate students
Premkumar et al. [7]	25	100	3	M = 35.4 years (SD = 11.9), 9 female
Steingroever et al. [8]	70	100	2	M = 24.9 years (SD = 5.8), 49 female
Steingroever et al. [9]	57	150	2	M = 19.9 years (SD = 2.7), 42 female
Wetzels et al. [15][c]	41	150	2	Students
Wood et al. [16]	153	100	3	M = 45.25 years (SD = 27.21)[d]
Worthy et al. [17]	35	100	1	Undergraduate students, 22 female

Introduction - Experimental Framework

- Payoff Scheme:

1) uses variable loss in deck C + has fixed wins and losses

2) uses constant loss in deck C + within each deck a random repetition of the payoff

3) uses variable losses and wins + has fixed wins and losses

   + changes rewards and losses → net outcome of bad decks becomes increasingly negative every 10 trials and vice versa

Procedure

- Experimental Framework
- Experimental Designs/ Sampling Strategies:
    - Full Factorial
    - Random Sampling
    - Active Learning:

      1) More risky choices but >threshold score in the end
      2) Less risky choices but >threshold score in the end
      3) More risky choices but <threshold score in the end
      4) Less risky choices but <threshold score in the end
- sklearn estimator:

  1. DecisionTreeRegressor
  2. RandomForestRegressor
  3. GradientBoostingRegressor
  4. KNeighborsRegressor

Results

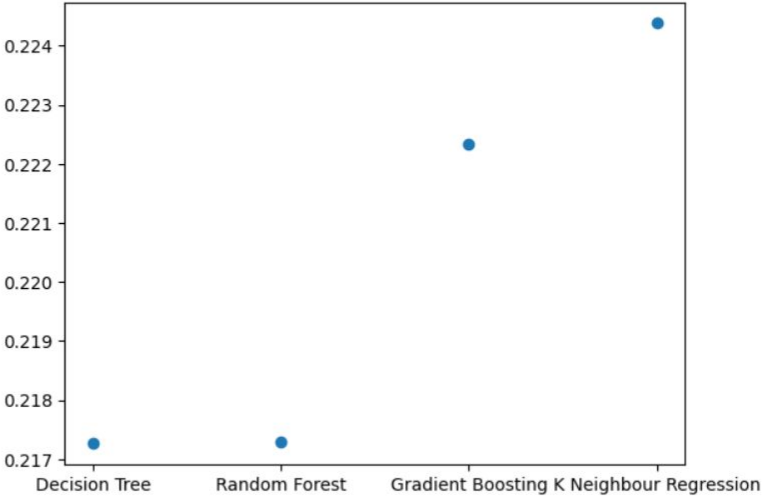
Estimators with the dependent variable ‘Choice’:

- **DecisionTreeRegressor: 0.8042019639192353**
- **RandomForestRegressor: 0.8041715941535984**
- **GradientBoostingRegressor: 0.7996251807103818**
- **KNeighborsRegressor: 0.797787051892027**

Interpretation - Comparison

Comparison of 4 different estimators

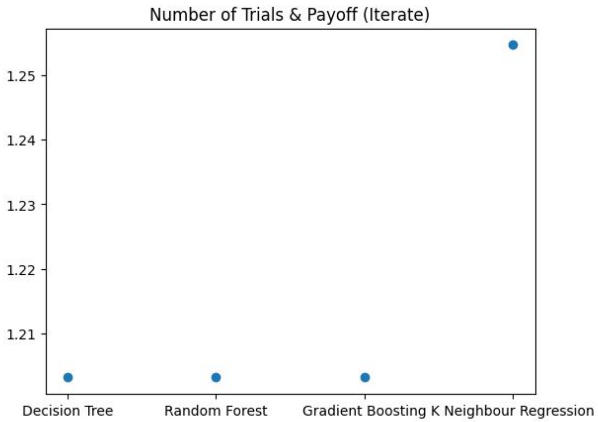
```
Estimator: Decision Tree
Mean Squared Error: 0.21726034944776285

Estimator: Random Forest
Mean Squared Error: 0.21728410819980187

Estimator: Gradient Boosting
Mean Squared Error: 0.22233881468268482

Estimator: K Neighbour Regression
Mean Squared Error: 0.22437842916196918
```

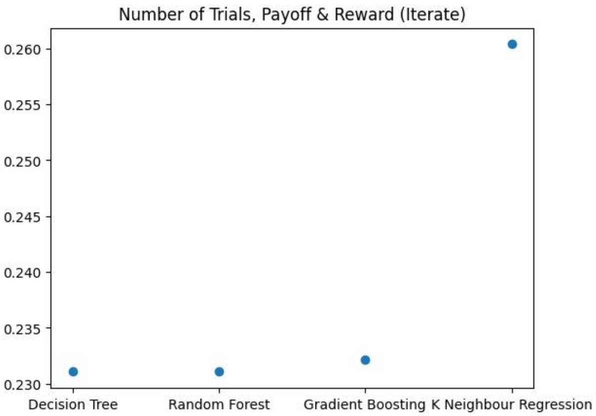


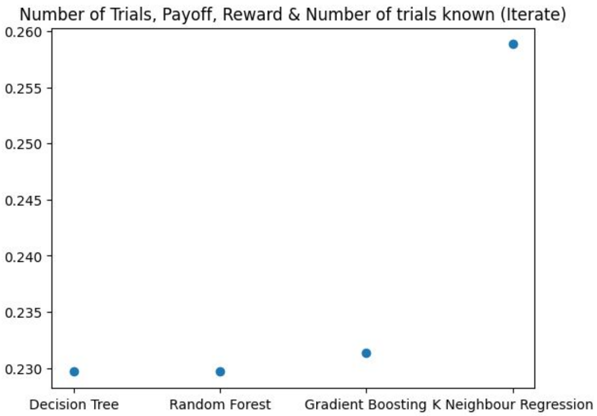
Full Factorial Design

- Set up experimental design using sweetpea
- Payoff & number of trials as factors -> otherwise no fitting samples
- Found samples according to synthesized experiment
- **Sampling strategies:** IterateGen, CMSGen

Full Factorial Design - Results

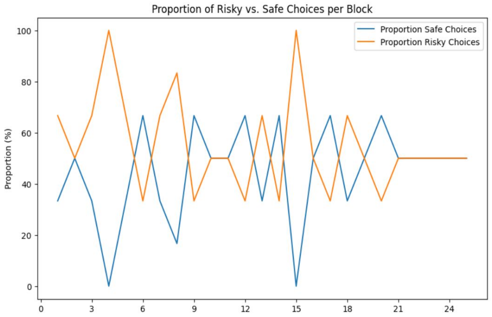


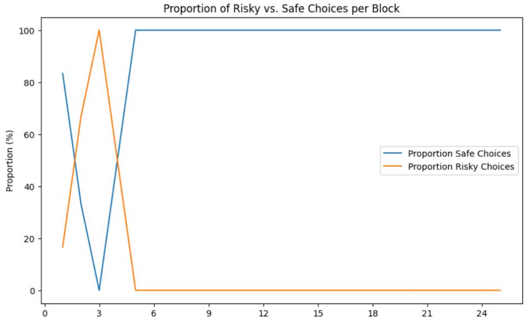


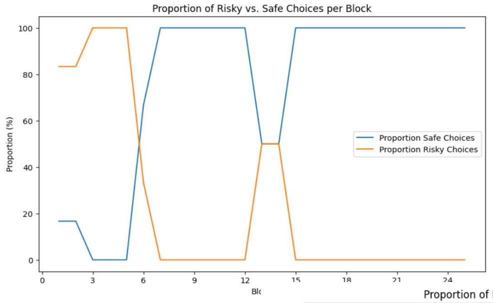


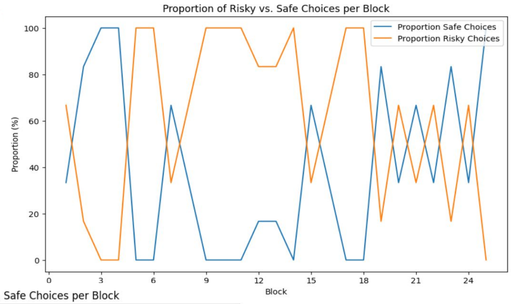
The initial variable pattern:

Active Learning:



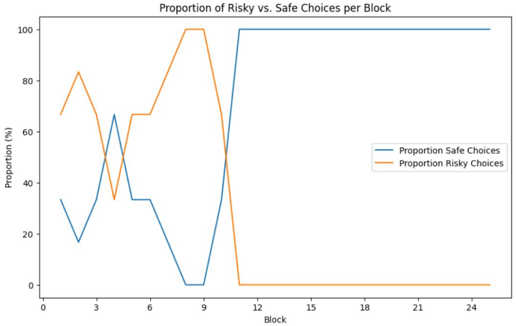






Proposition of Risky vs. Safe Choices

x-axis: Block
y-axis: Proportion (in %)



Anish Haluvani Sundresh, Hannah Köster, Jennifer Kempkens, Rowan Mohamed, Jessica Simon

Interpretation - Comparison

Active Learning:

- XGB Committee
- Random Forest Committee

+ Using the estimators directly on the entire training without AL and querying

- XBG Committee: rmse=21.41, mae=14.67
- Random Forest Committee: rmse=21.41, mae=14.67
- AL modeling passive_profit: rmse=21.41, mae=14.67

Conclusion

Did particular estimator outperform the others or did they perform equally well?

- **4 Estimators perform rather equally**

  → Decision Tree Regression performs best

  → Gradient is boosting slightly worse

- **In Full Factorial Design (sampled data)**

  → Decision Tree Regression and Random Forest perform best

- **Random Forest works better than XGB or Gaussian due to the types of variables.**