Iron Insight

Eddie Kiernan, Logan Proffitt, Noah Turner

Problem Space:

Lifters of all levels can often feel discouraged about their progress and wonder how much they might be able to lift in the long run. The presence of performance enhancing drugs (PEDs) has only added to this discouragement, and skewed people's perceptions of what is possible. We would like to give lifters a better idea of how much they might be able to lift with and without PEDs, based on the progression of competing powerlifters.

Dataset:

Our data set contains over 1,000,000 entries. Each entry contains 40 features, the most important being Name, Sex, BodyWeight, Date, Tested, Division and Best3Bench. The outcome variable would be a bench press increase over time. We will clean the data so that each entry is a lifter at a given meet who has had at least 1 meet and valid bench numbers at each meet.

Approach:

We plan to use a hybrid model for this problem, combining regression or tree models with other time-series models that consider time-dependent variations. We will begin by using a KNN or tree-based model to find lifters that were similar to the user at a given meet (in terms of age, bodyweight, sex, tested vs untested, current bench, goal bench). Then, we will use a time-series model to make a prediction based on the progression of those similar lifters.

Considerations:

We will likely bake in a cap on the model's output, since progression does not continue indefinitely (we don't want to tell someone that weights 135lbs that they will bench 700lbs).

We will also have to factor in whether a lifter is tested or not, and whether they may have switched back and forth between tested and untested throughout their career. The user may be able to compare possible future lifts based on natural lifters and untested lifters.

Stretch Goals

If time allows, it may be worth implementing a percentile system, where a user can see how realistic a given range of predictions might be. For example, if a prediction is based more heavily on world class athletes, it might be placed into a higher percentile to indicate that it could be less realistic. Likewise, if more lifters have achieved a lighter lift, the user would be able to tell that this is more achievable.

Another interesting feature would be addressing how predictions would change with PEDs. We could allow the user to see how much faster a lift might be reached with the use of PEDs, or how much more they could potentially lift.

Research Papers

One study (https://sportrxiv.org/index.php/server/preprint/view/218) looks at the quality of strength gained in different phases of lifting. They found that athletes typically made more gains early on, but less as time went on. They used similar data on powerlifters and focused on similar features to the ones we will be using.

Another study (https://pubmed.ncbi.nlm.nih.gov/35591809/) looked at long term strength development from minimal-dose resistance training. They found that participants made the best gains in strength during the first 1 to 2 years but plateaued soon after. The overall progression was logarithmic, increasing rapidly at first but seeing diminishing returns as time went on.