

## Iron Insight

Eddie Kiernan, Noah Turner, Logan Proffitt

### Problem Space

Today, the perceptions of lifters are more skewed than ever. Social media promotes the most elite of lifters, and the presence of performance enhancing drugs (PEDs) has made it increasingly difficult to tell what is possible naturally. As a result, many lifters have little idea of what is normal, and what progress is actually possible. This uncertainty often leads to discouragement, as lifters feel they aren't achieving what they should.

Our group aims to address this issue by providing accurate and realistic predictions for a natural lifter's future bench press. Our approach differs from others in the following ways:

1. Most lifting calculators focus on predicting 1 rep maxes, or telling you how good your lifts are. For example, strengthlevel.com (one of the most popular lifting websites) takes your age, bodyweight, reps, and weight, and gives a predicted 1 rep max and your percentile. These websites are very useful for planning workouts, but provide no insight into what your future progression might look like. Instead, our models predict an *increase* based on the lifter's data and a given amount of time.
2. Another difference is the data we are using. Many lifting websites (including strengthlevel.com) use user-entered data, which opens the door for falsified lifts and the use of PEDs. This isn't a big deal when predicting 1 rep maxes, but makes a huge difference when it comes to progression. To avoid this issue, our team is using competition data with a tested/untested feature, allowing us to filter out most PED users. This allows us to make predictions better suited for natural lifters.

### Approach

Our main approach was a KNN Regression model. The data used for this model was first grouped so each datapoint was an individual who had a list of valid bench presses and a number of days after start corresponding to each bench press.

## Project Report

	BestBench	Sex	StartBench	StartAge	StartWeight	DaysAfterStart
0	[130.0, 145.0]	M	130.0	27.5	72.50	[0, 184]
1	[165.0, 185.97, 155.0, 195.0, 195.0]	M	165.0	21.5	92.00	[0, 266, 441, 749, 749]
2	[170.0, 182.5]	M	170.0	22.5	116.40	[0, 182]
3	[65.0, 75.0]	F	65.0	22.5	62.40	[0, 364]
4	[100.0, 105.0, 112.5, 125.0]	M	100.0	18.5	71.50	[0, 180, 364, 878]
...	...	...	...	...	...	...
161540	[162.5, 182.5, 182.5]	M	162.5	20.5	122.95	[0, 714, 714]
161541	[95.0, 95.0]	M	95.0	19.5	61.55	[0, 183]
161542	[140.0, 135.0]	M	140.0	19.5	83.00	[0, 364]
161543	[120.0, 152.5, 160.0]	M	120.0	44.0	86.90	[0, 1294, 1609]
161544	[100.0, 95.0, 100.0, 100.0]	M	100.0	68.5	73.15	[0, 364, 728, 1078]

161531 rows x 6 columns

Our KNN model was fit by one hot encoding the 'Sex' feature, scaling the 'Bench', 'Bodyweight', and 'Age' features, and fitting a specified regression model for each datapoint on 'BestBench' and 'DaysAfterStart'. Predictions for a future bench press were done for a datapoint by inputting a starting 'Bench' value, 'Sex', 'Age', 'Bodyweight', and a number of days to predict. A BallTree tree was used to find the k nearest neighbors, then the neighbors' estimated bench for the input number of days was predicted using the that specified neighbors fitted regression model, and the results were averaged together for the prediction for the input data point. We experimented with different weighted averages, but found that a normal average performed best.

Other types of approaches we tried were plain Regression models and a Neural Net. The plain Lasso, Ridge, and ElasticNet regression models were constructed by adding 'DaysAfterStart' and 'BenchIncrease' features for each data point. Additionally, log transforms were applied to various features throughout testing. However, these models underperformed compared to the KNN model.

Further, the attempted Neural Net model achieved a respectable performance with an RMSE of 11.67 kg and a MAPE of 6.81% meaning that the predictions of future bench presses were within 6-7% of the actual values. While the model tried to capture overall trends the residual graph shows a negative trend, suggesting that the model overpredicts for higher values, additionally the farther out the model predicts the more the variance increases, these could be due to things that the model cannot take into consideration like plateaus, injuries, live events, etc. Short term predictions within ~5 years the model was notably more accurate, however the neural network still underperformed the KNN in these time frames and this is likely due to the absence of some key features such as 'YearsTrained' which could provide key insight into whether a lifter was new or already experienced. The model also searches for a global fit vs a local fit trying to generalize across all of the data rather than the variations in local

## Project Report

minimums. Despite the limitations of this model in our environment it provided us with a good baseline to improve upon for future models and techniques.

### Data

Our data came from OpenPowerlifting, an open source community service project to create a permanent, open archive of the world's powerlifting dataset. The bulk data set we used contains 3,290,035 entries, where each entry is a single individual competing in a specific meet on a specific date. Each entry contains 42 features however for the purposes of our project we only utilized the following key features: "Name", "Sex", "Age", "BodyweightKg", "WeightClassKg", "Best3BenchKg", "Date", "Tested", and "Equipment".

	Name	Sex	Event	...	MeetTown	MeetName	Sanctioned
0	E.S. Denisenko	F	B	...	Luninets	Bison Power Cup	Yes
1	I.S. Lebedskaya	F	B	...	Luninets	Bison Power Cup	Yes
2	K. Yakimovich	F	B	...	Luninets	Bison Power Cup	Yes
3	A.G. Golneva	F	B	...	Luninets	Bison Power Cup	Yes
4	E.V. Marunetskaya	F	B	...	Luninets	Bison Power Cup	Yes
...	...	...	...	...	...	...	...
3290030	Ibrahim Sandogji	M	SBD	...	Riyadh	The Kingdom Classic Powerlifting Championships	Yes
3290031	Feras Alwahhabi	M	SBD	...	Riyadh	The Kingdom Classic Powerlifting Championships	Yes
3290032	Musab Alessa	M	SBD	...	Riyadh	The Kingdom Classic Powerlifting Championships	Yes
3290033	Ali Alholan	M	SBD	...	Riyadh	The Kingdom Classic Powerlifting Championships	Yes
3290034	Alfarooq Eissa	M	SBD	...	Riyadh	The Kingdom Classic Powerlifting Championships	Yes

3290035 rows x 42 columns

After cleaning the data to only include lifters who had multiple occurrences in the data (multiple lifts over time), had been tested for performance enhanced drugs, only competed in with raw or wraps for their equipment, and removed any Null, NaN or invalid for other key features the cleaned data resulted in 726,426 rows and 7 features.

	Name	Sex	Age	BodyweightKg	WeightClassKg	BestBench	Date
0	Maysar Olaymi	M	24.0	43.5	49.0	91.0	2017-02-21
1	Mohammad Alshnaiti	M	15.0	39.5	49.0	70.0	2017-02-21
2	Ali Al-Darraj	M	35.0	52.3	54.0	142.0	2017-02-21
3	Ziad Androun	M	31.0	63.5	65.0	131.0	2017-02-21
4	Hasan Al-Tameemi	M	39.0	64.4	65.0	157.0	2017-02-21
...	...	...	...	...	...	...	...
726477	Mostafa Ali Alfaky	M	35.0	146.3	120.0	155.0	2022-10-13
726478	Ibrahim Sandogji	M	17.0	64.4	66.0	120.0	2022-10-13
726479	Feras Alwahhabi	M	17.0	82.9	83.0	112.5	2022-10-13
726480	Musab Alessa	M	17.0	91.3	93.0	130.0	2022-10-13
726481	Alfarooq Eissa	M	18.0	157.3	120.0	130.0	2022-10-13

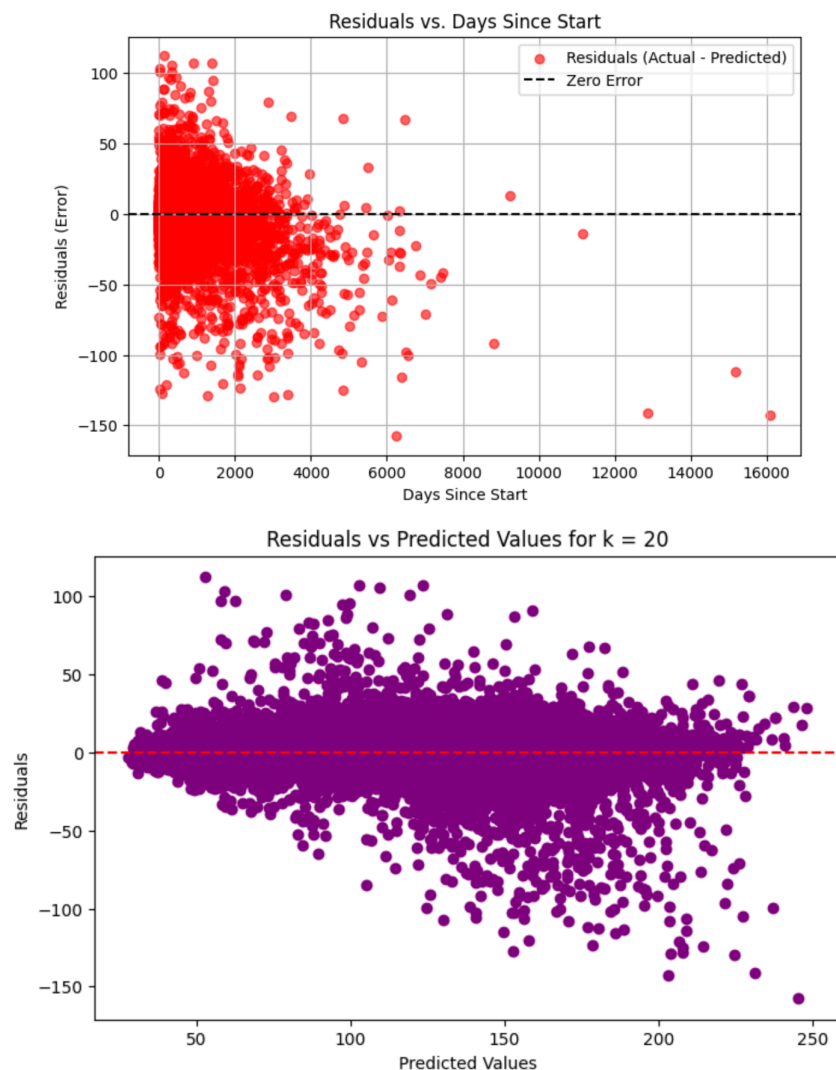
726426 rows x 7 columns

## Project Report

The outcome variable we intend to predict is future bench press given a number of days, from a start point at a current bodyweight, sex, age, and current bench press. Further data cleaning and feature engineering was carried out for the purposes of each specific model.

### Results

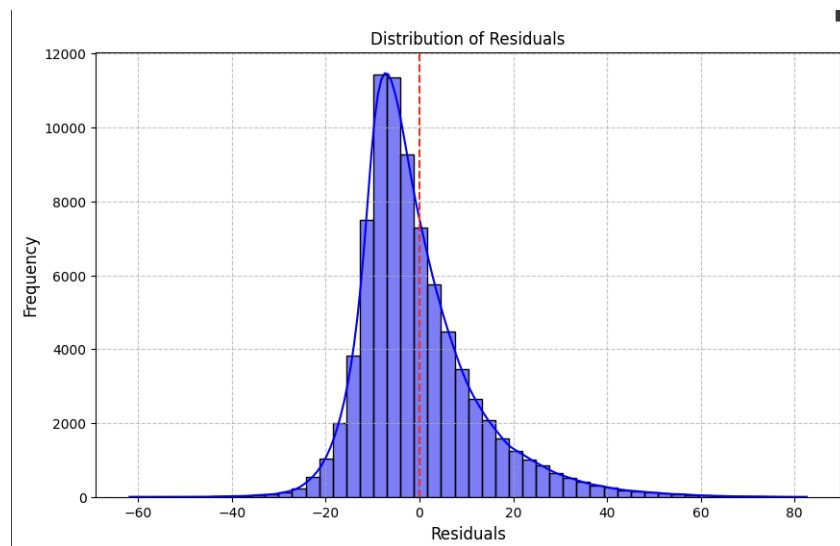
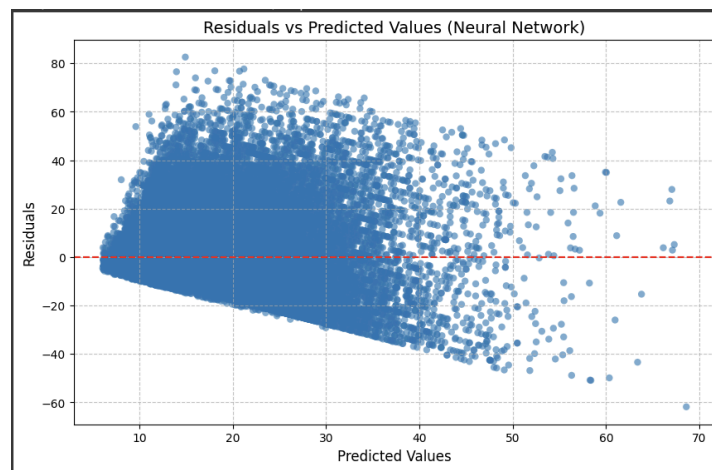
For our KNN model testing was done to find the best regression model and which k-value, it was determined a k value in the range of 20-30 yielded the best results with an ElasticNet regression model with alpha and l1 ratio parameters of 10,000 and .01 respectively. The best error calculations achieved were an RMSE of 11.6756, Mean Absolute Percentage Error (MAPE) of 6.81% and an Average Error of 6.6611. Below are the calculated residual graphs for the KNN model.



## Project Report

We also tried plain ridge, lasso, and elastic net regression models. While these models had an RMSE of 11.36 kg, they significantly underestimated long term progression, performing worse than the KNN overall. This poor performance could likely be attributed to training on the dataset as a whole, and failing to capture the variance between different lifters.

For our neural network model, we tested how well it could predict bench press increases. The model had a Root Mean Square Error of about 11.67 kg and a Mean Absolute Percentage Error of 6.81%, meaning the predictions were off by about 6-7% on average. The Residuals vs Predicted Values graph shows that the model tends to overpredict higher bench press values, with the errors getting bigger as the predictions increase. The Distribution of Residuals graph shows most errors are close to zero, but there is a slight positive bias, meaning the model often predicts values that are a little too high. Overall, the model works well for most predictions but struggles a bit with bigger bench press increases, especially for more advanced lifters. Adding features like Years Trained or training details might help improve the model in the future.



### Discussion

The results from the KNN Regression model are telling of a few key things. First, the measures of error, specifically the ~6 Average Error, mean when predicting on new data the model is roughly 6kg off for a given bench press. This result is respectable as for the amount of variance present among lifters, whether its programming, diet, recovery, or genetics, being able to predict future bench progression within  $\pm 6$ kg exceeded our initial expectations. Second, from the Residuals vs Days Since Start graph the model becomes more unreliable in its predictions after about 4000 days (~10 years), and gives best predictions within the scope of 2000 days (~5 years). This was expected before beginning this project as the element of time, whether due to injuries, falling out of love with the activity of weightlifting, or general aging can all lead to long-term plateaus and resulted in not having a good amount of long term representation in our data. Finally, the Residuals vs Predicted Value graphs have a distribution that is mostly randomly scattered, but there is a somewhat negative linear trend. This negative trend is an indication of slight over predictions by the model. The most likely explanation for this is not having a feature to tell us how long a lifter has been lifting for. The current scientific literature on strength training indicates that new lifters can progress at a much faster rate than experienced lifters. While our data tells us the first occurrences when someone competes in a powerlifting meet, it does not indicate how long they have been training for. So, one individual's first meet could come after 10 years of training, meaning the ceiling for future progress is likely lower and would take a lot more time, while a different individual's first meet could come after 6 months of training, giving them a much higher ceiling and faster progression.

In totality, the results achieved by the KNN Regression model were satisfactory, but could be improved with further investigation into certain additional features to the data. Additional features that could improve the model in the future would include 'YearsTrained' as discussed prior, features like 'TotalAverageSleep', 'Diet', 'TrainingProgram' to capture specific types of training, diet, and recover, and features that can capture genetic advantages or disadvantages like 'FreeTestosterone' and 'GenomeExpression'. Additional features could improve upon the general accuracy of the model's performance and could also potentially lead to a better understanding of what specific features are most crucial to predict strength gains. For example, what's more important for predicting strength gains, 'Age', 'Bodyweight', 'Diet', or something like 'GenomeExpression'? These findings could help give lifters a clearer picture of what to prioritize when it comes to strength training.

Future plans for this problem could include finding datasets with a YearsTrained feature or other more specific features (diet, sleep, etc.) that could improve predictions. We could also experiment with other models such as clustering, or other means of aggregating the regressions of neighbors.