

Topic Project 4.1

# Applying Supervised Learning to Predict Student Dropout

Diptarko Chowdhury  
26<sup>th</sup> November 2024

The PureGym Group has over 2 million members and 600 gyms across the world. To ensure continued success, maintaining a high level of customer satisfaction is key. The reviews left by PureGym users can be analysed to inform stakeholders as to the group's successes and failures, as well as advise next steps when it comes to improving the value proposition that PureGym is built upon. By being responsive to customer concerns and addressing shortcomings, PureGym hopes to retain existing members, gain new members and increase the overall business value.

Two datasets of customer reviews are used, `Trustpilot_12_months.xlsx` and `Google_12_months.xlsx`. Across both datasets there are 39,923 reviews on 579 locations. Both datasets contain three key columns which are explored in this analysis, which contain:

- To get a general understanding of the reviews, the Natural Language Toolkit (NLTK) was used on the reviews (after they had been pre-processed and tokenized), to reveal the most common words across both datasets as can be seen in Figure 1.

[illegible]

Figure 2: Word clouds for reviews in both datasets





Figure 5: The top topics and contributing words for all negative reviews, as identified by BERTopic

From topics 0 – 4 as labelled we can infer that common complaints are around

- Ineffective air conditioning resulting in gym that is too hot
- Rude PureGym staff
- Unsatisfactory maintenance of changing rooms and toilets
- Irritating parking protocol
- Odorous environment because of poor ventilation

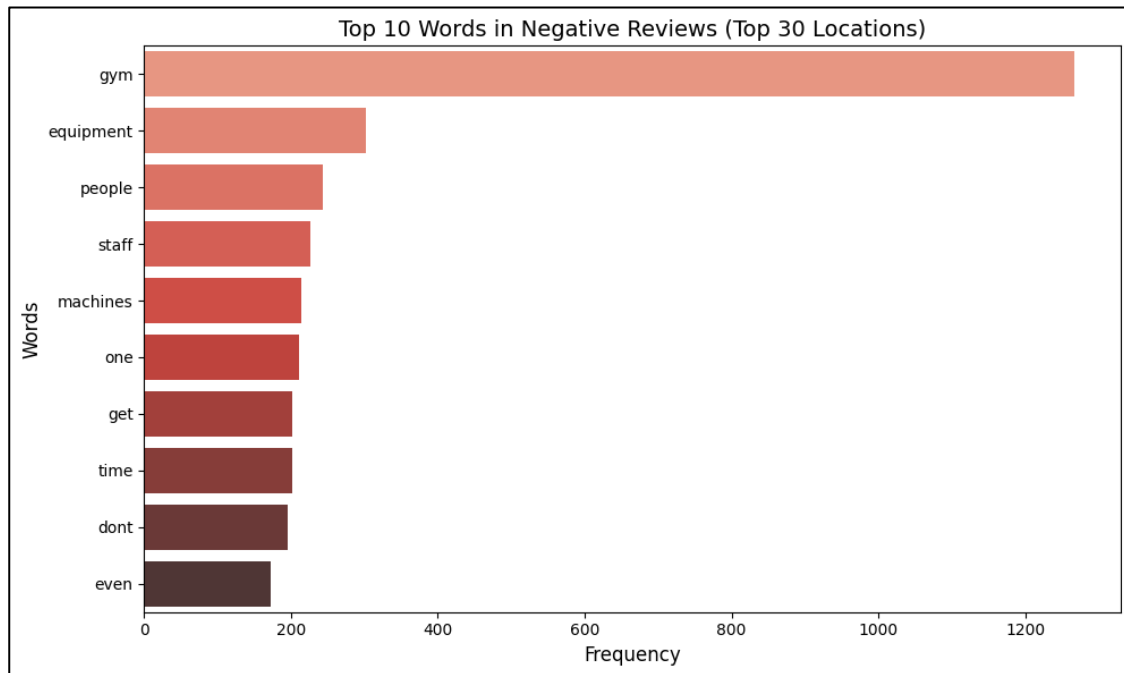
## Identifying Poorly Reviews Locations

For further analysis the 20 PureGym branches with the most negative reviews were extracted.

Top 20 Locations with Negative Reviews (Google):			Top 20 Locations with Negative Reviews (Trustpilot):		
Location	Negative_Review_Count		Location	Negative_Review_Count	
319 London Stratford	59	155 Leicester Walnut Street	50		
275 London Canary Wharf	26	0	345		
330 London Woolwich	26	186	London Enfield	23	
285 London Enfield	25	219	London Stratford	22	
321 London Swiss Cottage	24	58	Burnham	20	
311 London Palmers Green	22	199	London Ilford	18	
124 Birmingham City Centre	21	171	London Bermondsey	18	
303 London Leytonstone	21	233	Maidenhead	16	
353 New Barnet	20	215	London Seven Sisters	16	
380 Peterborough Serpentine	19	195	London Hayes	16	
444 Wakefield	19	188	London Finchley	16	
137 Bradford Thornbury	19	358	York	16	
297 London Hoxton	18	261	Northwich	15	
315 London Seven Sisters	18	221	London Swiss Cottage	15	
445 Walsall Crown Wharf	18	174	London Bromley	15	
294 London Hayes	17	194	London Hammersmith Palais	15	
341 Manchester Exchange Quay	17	95	Dudley Tipton	14	
104 Bachembülach	17	44	Bradford Thornbury	14	
269 London Bermondsey	16	343	Watford Waterfields	14	
370 Nottingham Colwick	16	330	Telford	14	

Figure 6: The 20 branches with the most negative reviews

Branches appearing in both datasets should be investigated further. Notably, London Stratford and London Enfield appear in the worst 5 locations for both datasets. Taking the 30 locations with the most joint negative reviews, NLTK is used once again to get results as in Figures 7 and 8.



Perhaps predictably, there's not much different between these results and those seen in Figures 3 and 4. However when we feed only reviews from the 30 most negative reviewed locations into BERTopic, we do extract additional information.

- Users having difficulty with their membership/using pins for day passes
- Gyms being closed during advertised operational hours

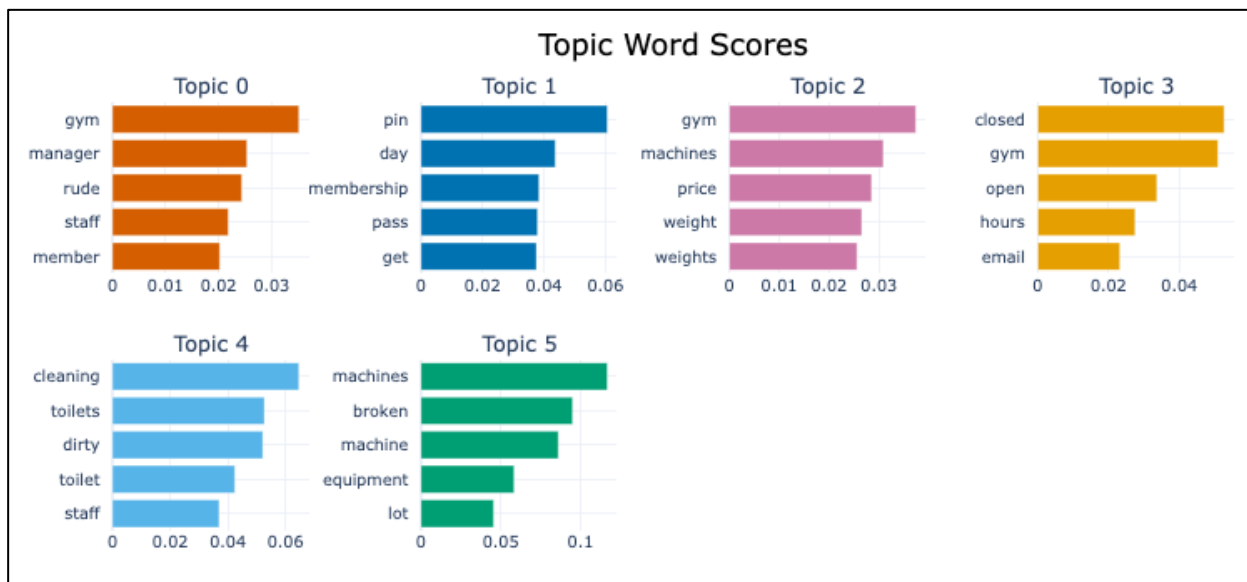


Figure 9: The top topics and contributing words for negative reviews on the 30 most negatively reviewed locations

### Emotion Analysis using a BERT Model from Hugging Face

To determine the underlying user emotion from these negative reviews, the BERT model `bhadresh-savani/bert-base-uncased-emotion` was deployed.

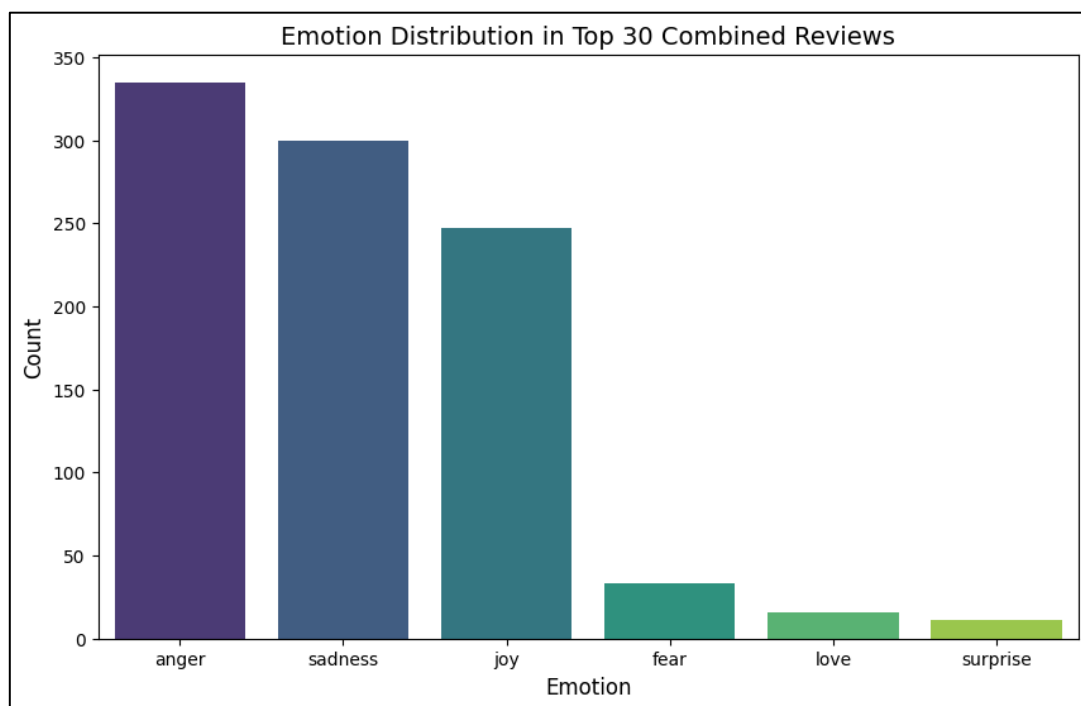


Figure 10: Emotion Distribution in reviews for the 30 most negatively reviewed locations

As seen in Figure 10, the presiding user emotion is anger, followed closely by sadness. The model's limitations are also apparent as joy and love rank third and fifth respectively. This suggests the model is struggling with complete sentiment analysis, perhaps around sarcasm or some of the oxymoronic idiosyncrasies particularly prevalent in British English.

Since anger is the strongest negative emotion, all predominantly angry reviews from this subset were taken and fed once more into BERTopic.

Topic 0: equipment, gym, weights, machines, people, machine, bench, space, use, one

Topic 1: er, og, ikke, det, der, til, de, af, en, jeg

Topic 2: pass, day, code, pin, bought, didnt, paid, get, work, let

Topic 3: rude, member, manager, staff, gym, personal, aggressive, attitude, pt, trainer

Topic 4: und, die, nicht, ich, das, ist, es, zu, sehr, sind

Topic 5: class, classes, booked, cancelled, book, instructors, instructor, one, week, pilates

Topic 6: membership, cancel, account, month, suspended, cancelled, payment, email, cancellation, charged

Topic 7: fee, joining, code, promo, discount, charged, applied, website, offer, membership

Topic 8: membership, joining, fee, access, gym, gyms, pay, month, paid, open

Topic 9: toilets, toilet, cleaning, changing, disgusting, room, cleaned, rooms, sewage, towel

Topic 10: cleaning, disgusting, sweat, gym, stations, smell, air, smelly, equipment, spray

Topic 11: parking, car, park, fine, fines, free, reg, fined, appeal, received

Topic 12: rude, staff, manager, unhelpful, extremely, tasleem, unprofessional, customers, disrespectful, attitude

Topic 13: music, loud, classes, hear, noise, play, headphones, extremely, gym, shouting

Topic 14: people, gym, around, members, weights, staff, equipment, kids, teenagers, school

Topic 15: cold, showers, shower, hot, water, temperature, gym, changing, warm, rooms

Topic 16: closed, open, christmas, gym, closing, day, days, hours, opening, due

Topic 17: cold, showers, water, freezing, shower, warm, boiler, every, months, issue

Topic 18: air, ac, conditioning, summer, hot, airconditioning, unbearable, aircon, sweat, con

Topic 19: pin, sent, didnt, log, email, app, tried, give, send, given

Topic 20: lockers, stolen, locker, lock, stole, bag, security, room, safe, thieves

Figure 11: The top 20 topics and contributing words within angry reviews on the 30 most negatively reviewed locations

Figure 11 flags further issues to take note of such as:

- Pre-booked classes being cancelled at short notice
- Difficulty joining the gym
- Noise pollution
- Cold showers
- Theft

## Using a Large Language Model from Hugging Face

The `tiiuae/falcon-7b-instruct` model was used on all negative reviews with the prompt “Below is a customer review. Identify the main 3 topics related to customer experience. These topics were then collated to create a comprehensive topics list, which a section of can be seen in Figure 12.

```
[ 'Poor customer service',
  'Inadequate facilities',
  'Unfriendly staff',
  'Staff attitude and communication',
  'Cleanliness and facilities',
  'Variety of equipment and classes',
  'Cleanliness',
  'Staff Knowledge',
  'Timeliness',
  'Dirty machines',
  'Poor hand hygiene',
  'Inadequate cleaning protocols',
  'Poor equipment and cleanliness',
  'Unhelpful and unprofessional staff',
  'Dirty lockers',
  'Gym cleanliness',
  'Customer service',
  'Maintenance',
```

Figure 12: A selection of the topics produced by the LLM



This list was then fed into another `tiiuae/falcon-7b-instruct` model with the prompt “For the following text topics obtained from negative customer reviews, can you give some actionable insights that would help this gym company?”. A selection of suggestions feature in Figure 13.

Actionable Insights:  
Implement a customer service training program to improve staff's ability to handle customer complaints and provide better overall service.  
Regularly clean and maintain equipment to ensure a safe and hygienic environment for members.  
Invest in modernizing and updating equipment to keep up with industry standards and provide a better workout experience.  
Offer incentives or rewards to encourage members to maintain their commitment and promote word-of-mouth marketing.  
Address the inadequate facilities issue by investing in proper equipment, hiring experienced staff, and providing regular maintenance to ensure top-notch facilities.  
Implement customer service training to ensure a high level of customer satisfaction and retention.  
Consider revamping the pricing strategy by offering flexible packages, student discounts, or loyalty programs to attract and retain customers.  
Train staff in customer service and communication skills to improve customer experience.  
Implement strict cleanliness protocols and regularly maintain equipment to ensure a safe and hygienic environment.  
Offer incentives or rewards to attract and retain customers, and actively address and resolve customer complaints to build a positive reputation.  
Implement a staff training program to improve customer service and communication skills.  
Perform regular equipment maintenance to ensure proper functioning and reduce breakdowns.  
Establish clear guidelines for staff on how to handle customer complaints and feedback.  
Enhance the gym's cleanliness by increasing the frequency of cleaning and disinfecting equipment and surfaces.  
Offer incentives or rewards to encourage staff to maintain a high level of customer satisfaction.  
Implement a rigorous cleaning schedule to maintain a clean and hygienic environment for members.  
Train staff to provide exceptional customer service and attentiveness to members.  
Consider offering flexible pricing plans or discounts to attract new customers and retain existing ones.  
Gym should focus on providing a variety of equipment and classes to cater to different fitness needs and interests. This can help attract and retain a diverse range of members.  
Improving customer service by hiring and training staff to provide personalized attention and support can go a long way in building a loyal customer base.  
Implementing cleanliness protocols and maintaining a clean environment is crucial for customer satisfaction and brand reputation.

Figure 13: A selection of the suggestions given by the LLM

Conclusion

PureGym reviews are largely glowing, but to improve overall customer experience, negative customer reviews should be analysed and subsequently suggestions for improvement followed. 15 summary suggestions using the actionable insights, made by inputting the full Figure 13 results into a GPT model feature in Figure 14.

1. **Enhance Customer Service Training:** Train staff to improve communication, empathy, and active listening skills to handle complaints effectively and enhance the overall customer experience.

2. **Implement a Rigorous Cleaning Schedule:** Regularly clean and sanitize gym facilities, including equipment, changing rooms, and common areas, to maintain a hygienic environment.

3. **Modernize and Maintain Equipment:** Invest in upgrading outdated equipment, perform regular maintenance, and replace broken machines to ensure safety and efficiency.

4. **Improve Air Conditioning and Ventilation:** Install or upgrade air conditioning systems and maintain proper ventilation to enhance comfort for members during workouts.

5. **Introduce Customer Feedback Systems:** Collect, analyze, and act on customer feedback to identify pain points and implement necessary improvements.

6. **Enhance Gym Layout and Space Management:** Reorganize the layout to reduce overcrowding, improve accessibility, and create designated areas for different activities.

7. **Offer Flexible Membership Plans:** Provide affordable and flexible pricing options, such as discounts, loyalty rewards, and tailored packages, to attract a broader audience.

8. **Increase Staff Presence During Peak Hours:** Hire additional staff to manage busy periods, ensuring better service and quicker response times.

9. **Implement Equipment Reservation Systems:** Introduce systems that allow members to reserve equipment or classes, minimizing wait times and overcrowding.

10. **Provide Additional Amenities:** Offer services like towel provision, water fountains, locker ventilation, and improved shower facilities to enhance the member experience.

11. **Ensure Safety Protocols Are in Place:** Conduct regular safety checks, train staff on emergency response, and maintain equipment to prevent injuries and ensure member safety.

12. **Improve Marketing and Communication:** Use social media and digital campaigns to highlight improvements, respond to feedback, and engage with both existing and potential members.

13. **Address Staff Attitude Issues:** Encourage professionalism, friendliness, and attentiveness among staff through incentives, training, and clear policies.

14. **Expand Class Variety and Scheduling:** Offer diverse classes catering to different fitness levels and preferences, with flexible timings to suit a wider audience.

15. **Maintain Transparency with Members:** Communicate changes, updates, and closures proactively to avoid frustration and build trust with members.

Figure 14: 15 Summary suggestions for PureGym to address negative reviews