CE807-7-SU: Text Analytics - Review Rating Classification Report

Name: Thuan Anh Bui
Student ID: 2412204
Username: tb24600
Lecturer: Dr. Ravi Shekhar

Abstract

The project explores the impact of tokenization methods on multi-class text classification for predicting review ratings from 1 to 5. Two approaches were evaluated: traditional tokenization with spaCy lemmatization and TF-IDF, and the BERT tokenizer using subword-level encoding. Both were paired with a Logistic Regression classifier to isolate the effect of tokenization. Models were trained and tested on a dataset split into training, validation, and test sets. The BERT-based method achieved slightly better results, with an accuracy of 0.7271 and F1-score of 0.6445, compared to 0.7157 and 0.6316 for the traditional approach. The results suggest that modern tokenization offers performance gains even without full transformer-based models.

Materials

- Google Colab Code

- Zoom Recorded Presentation

- Google Drive Folder

1 Tokenization, Representation Discussion

1.1 Summary of 2 Tokenization

1.1.1 Traditional Tokenization

A custom preprocessing pipeline combined with a TF-IDF vectorizer and enhanced with lemmatization via spaCy, is utilized to transform raw review texts into normalized and sparse vector representations suitable for a Logistic Regression classifier.

Before the tokenization and lemmatization stage, the input text is first cleaned using the custom preprocessing function. It is then tokenized using the base analyzer from `TfidfVectorizer`. Tokens are rejoined and passed through spaCy's English model (`en_core_web_sm`) to extract lemmatized forms. This process enhances generalization and reduces dimensionality by grouping morphologically similar forms.

Both configurations are used to capture short multi-word expressions while controlling feature dimensionality. Given a raw text,

> exactly what I needed. works perfectly. Arrived on time.
Thank you

The following step-by-step procedure for tokenization and lemmatization:

- **Tokenization** - The string is split into tokens

```
['exactly',         'what',
'i',   'needed',   'works',
'perfectly', 'arrived', 'on',
'time', 'thank', 'you']
```

- **Lemmatization** - Each word is lemmatized using spaCy.

```
['exactly', 'what', 'need',
'work',        'perfectly',
'arrive',   'on',    'time',
'thank', 'you']
```

- **Stop Word Removal** - Common stop words are removed using the built-in English stopword list.

```
['exactly', 'need', 'work',
'perfectly',       'arrive',
'time', 'thank']
```

1.1.2 BERT Tokenizer

This approach leverages the subword-based WordPiece tokenizer from the `bert-base-uncased` model, as implemented in Hugging Face's `transformers` library. The tokenizer is used to segment text into context-sensitive subword units prior to vectorization using TF-IDF.

- **BERT Tokenization:** Each text is tokenized using `BertTokenizer`, which applies WordPiece segmentation. Tokens may include subword prefixes (e.g., "##able").

- **Post-processing:** The resulting subword tokens are joined with whitespace and passed to a `TfidfVectorizer` for representation.

These token vectors are then passed to a Logistic Regression classifier, forming the final pipeline. Using the same review text for the BERT Tokenizer,

- **BERT Tokenization** - The sentence is tokenized into subword tokens using `bert-base-uncased`:

```
['exactly', 'what', 'i',
'needed', '.', 'works',
'perfectly', '.', 'arrived',
'on', 'time', '.', 'thank',
'you']
```

- **Joined Tokens for TF-IDF Input** - Tokens are joined into a space-separated string:

  exactly what i needed . works perfectly . arrived on time . thank you

1.2 Critical discussion and justification

Both methods begin with an identical `preprocessing_stage`, ensuring a consistent and standardized input for the subsequent tokenization phase.

Traditional Pipeline - Using a `TfidfVectorizer` in combination with the spaCy library for lemmatization. Its key characteristics include:

- **Lemmatization:** Words are reduced to their base or dictionary form (e.g., "works" becomes "work," "needed" becomes "need").

- **Punctuation Removal:** The default tokenizer discards punctuation, treating it as noise.

- **Stop Word Removal:** Common English stop words (e.g., "what," "on," "you") are filtered out due to their low semantic value.

This approach heavily simplifies the text, focusing on the core meaning of the most significant words.

BERT-Based Pipeline - Leveraging a pre-trained `BertTokenizer` before feature calculation. Its approach is fundamentally different:

- **WordPiece Tokenization:** A subword-level tokenization strategy is used, allowing words to be split into meaningful fragments.

- **Punctuation Retention:** Punctuation is preserved as separate, meaningful tokens to aid contextual understanding.

- **No Lemmatization or Stop Word Removal:** All original words and tokens are retained exactly as produced by the tokenizer.

This pipeline preserves much more of the original grammatical structure and textual detail.

Empirically, we expect the BERT-based tokenization to outperform the traditional method, especially in handling rare words and in capturing subtle distinctions between ratings. However, it also increases preprocessing complexity and token sequence length, which can impact training time and memory usage during TF-IDF vectorization.

2 Design, implementation of classifiers

2.1 Data

The dataset used for this task comprises labeled user reviews for rating prediction, where the rating is an integer between 1 (lowest) and 5 (highest). Each sample contains the following columns: id (unique identifier), $data\_id$ (student-specific identifier), $text$ (review content) and $rating$ (target label). With an exception to the test dataset, there is no rating label column and there are two extra fields, namely $out\_label\_model\_1$ and $out\_label\_model\_2$. These are used to record predicted ratings from models trained using the two tokenization techniques—custom preprocessing and BERT tokenizer, respectively.

Split	# Samples	Columns Included
Train	4,914	`id`, `data_id`, `text`, `rating`
Validation	700	`id`, `data_id`, `text`, `rating`
Test	1,386	`id`, `data_id`, `text`, `out_label_model_1`, `out_label_model_2`

Table 1: Dataset Summary

In Figure 1, a higher allocation for a test set is needed to assess the effectiveness and performance of the trained model for each tokenization technique. Validation set is used mainly to assess the

accuracy rates after training but cannot be used for final assessment.
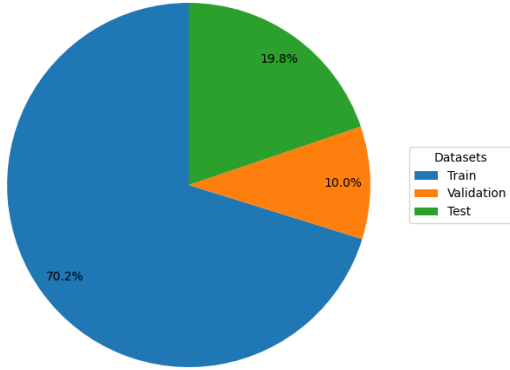

Figure 1: Dataset Percentage Composition

Figures 4 provide a breakdown of the number of reviews and their average word lengths per rating label in the training and validation sets. Both sets are highly imbalanced, with rating 5 making up over 66% of the training set and 66% of the validation set. This class imbalance poses a challenge for fair model evaluation. Notably, lower-rated reviews tend to be longer, while higher-rated ones (especially rating 5) are shorter on average. This may reflect greater elaboration in negative feedback.

	Rating	Number of Reviews	Average Length (Words)
0	1	669	45.50
1	2	218	53.27
2	3	277	49.17
3	4	501	55.16
4	5	3249	33.72
5	TOTAL	4914	nan

Figure 2: Review Dataset Summary in Training Set

	Rating	Number of Reviews	Average Length (Words)
0	1	95	51.34
1	2	31	44.84
2	3	39	50.62
3	4	72	45.69
4	5	463	39.01
5	TOTAL	700	nan

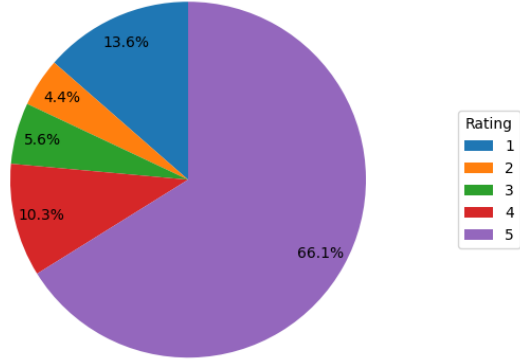Figure 3: Review Dataset Summary in Validation Set

Figure 4: Rating Distribution in Training and Validation Set

2.2 Model Implementation

The text preprocessing stage consists of the following steps:

- **HTML decoding**: Converts HTML entities into plain text (e.g., & → &).

- **Tag handling**: Replaces <br> tags with newline characters to retain formatting.

- **Punctuation normalization**: Collapses repeated punctuation (e.g., !!! → !) and replaces multiple hyphens with em dashes.

- **Whitespace cleanup**: Removes irregular spacing around punctuation and deduplicates whitespace.

- **Lowercasing**: Converts all text to lowercase to reduce vocabulary sparsity.

The following example of how a raw review is transformed step-by-step:

- **Initial Input: Raw Text**
  exactly what I needed. works perfectly. Arrived on time.
Thank you

- **HTML Tag Handling** - The <br /> tag is replaced with a newline character (\n).
  exactly what I needed. works perfectly. Arrived on time.\n Thank you

- **Normalize Punctuation and Spacing** - Extra spaces and redundant punctuation are removed. The newline is replaced with a single space.
  exactly what I needed. works perfectly. Arrived on time. Thank you

- **Lowercasing** - Converts all text to lowercase.
  exactly what i needed. works perfectly. arrived on time. thank you

2.2.1 Traditional Tokenization

The lemmatized tokens are vectorized using TF-IDF weighting. We experiment with two N-gram ranges: {unigrams, bigrams} and {unigrams, bigrams, trigrams}.

2.2.2 BERT Tokenization

The vectorizer is configured with: N-gram range between 1 and 3, maximum features of 20,000, sublinear TF scaling for term weighting.

2.3 Results

Model	Accuracy	F1 Score
Traditional Tokenizer	71.57	6.316
BERT Tokenizer	72.71	6.445

Table 2: Performance comparison between Traditional and BERT-based Tokenizers

3 Output Comparison and Analysis

3.1 Justification of model's performance

As shown in Figure 2, the BERT tokenizer outperformed the traditional TF-IDF-based method, achieving higher accuracy and F1-score. This suggests that BERT's context-aware subword representation enables better classification performance, especially for review texts with complex or nuanced language. The marginal improvement also reflects the efficiency of BERT embeddings in capturing semantic information that traditional bag-of-words approaches may overlook.

In the experiment, a BERT tokenizer in combination with a logistic regression classifier is employed to predict review ratings on a 5-point scale, achieving an accuracy of 72.71%. This demonstrates that even lightweight pipelines can yield reasonably strong performance using modern tokenization techniques.

A related study by (Biswas and Azharul Hasan, 2024) applied BERT tokenization in conjunction with a linear regression model on a significantly larger dataset of 112,000 reviews and reported 97.5% accuracy. Although the tasks appear similar, differences in modeling approach, dataset scale, and evaluation methodology (e.g., use of regression vs. classification) make direct comparison difficult. Nonetheless, the findings in both studies reinforce the value of BERT-based tokenization as an effec-tive foundation for downstream sentiment or rating prediction tasks.

3.2 Example and other analysis

To better evaluate model behavior, selected validation examples with varying sentiment and structural complexity were analyzed. Each case illustrates how specific linguistic features influenced correct or incorrect predictions. According to Table 3,

- **ID: 302654** A complex review mixing technical details and subjective praise. Both models correctly predicted the rating, showing robustness to mixed sentiment.

- **ID: 56099** This review shifts from neutral to strongly negative (e.g., *"I HATE THIS PIECE OF GARBAGE"*). Model 1 succeeded, likely due to sentiment keywords. Model 2 misclassified, possibly influenced by the neutral opening.

- **ID: 503538** A long, consistently positive review. Both models performed well; Model 1 responded to explicit praise, while Model 2 may have benefited from context retention.

- **ID: 574685** Begins with product marketing text, followed by light personal commentary. Both models misclassified it as overly positive, likely misled by the promotional tone.

- **ID: 136885** Structured like a manual, with scattered emotional remarks. Despite positive statements (*"I'm so glad I bought this washer"*), both models failed, likely due to dilution by technical content.

4 Discussion and Summary

These examples highlight that while both models can handle straightforward sentiment, they struggle with structural noise, indirect tone, or sentiment buried in lengthy explanations. BERT benefits from contextual depth but can misinterpret formatting or quoted passages, while traditional tokenization is more literal and sensitive to explicit sentiment cues.

4.1 Justification of Feature-based Algorithm

A critical focus was placed on the token-level representation of the input reviews, particularly through the use of **unigrams**, **bigrams**, and **trigrams**. To make this decision evidence-based, a preliminary

4

ID	Text (Full text in Appendix A)	True Rating	Model 1	Model 2
302654	I read several of the reviews…[R1]	5	5	5
56099	Didn't purchase from Amazon but posting…[R2]	1	1	5
503538	I'll be honest, I was very…[R3]	5	5	5
574685	What Amazon's product page says: \br / \br…[R4]	4	5	5
136885	[[VIDEOID:4fd222938153f33c6f93079d83e0720d]] I'm so glad I bought…[R5]	5	4	4

Table 3: Model predictions compared to true ratings. Green = correct, red = incorrect.

linguistic analysis of the review texts was conducted.

Adjective	Frequency
great	889
perfect	304
happy	170
nice	131
clean	106
better	137
best	97
excellent	81

Table 4: Frequencies of the most common adjectives for positive reviews (rating ≥ 4)

Table 4 shows that **unigrams** like *happy*, *nice*, and *perfect* are highly associated with positive sentiment. However, **unigram**-based models are limited in their ability to capture contextual negation or semantic shifts. For instance, the phrase *not happy*—despite containing the positive word happy—actually signals dissatisfaction.

Adjective/Phrase	Frequency
disappointed	33
expensive	21
very disappointed	21
not worth	18
too much	10
not happy	4

Table 5: Frequencies of the most common adjectives/phrases for negative reviews (rating ≤ 2)

As shown in Table 5, negated phrases such as *not happy* or *not worth* are strong indicators of negative sentiment, which would be missed by a purely **unigram**-based approach. This supports the integration of **bigrams** and **trigrams** into the model pipeline, enabling it to capture multi-word expressions and sentiment shifts that occur due to negation or intensification (e.g., very disappointed).

According to (Pranckevičius and Marcinkevičius, 2016), the level fo accuracy improves significantly from 41.27% to 78% if 1-2-3-grams are employed.

4.2 Comparison to State-of-the-Art Approaches

From an algorithmic perspective, Logistic Regression proved effective when paired with good feature selection and proper preprocessing. However, future work could explore more advanced models like fine-tuned transformer-based classifiers to handle deeper semantic understanding.

State-of-the-art models such as fine-tuned transformers (e.g., BERT, RoBERTa, DistilBERT) in (Joshy and Sundar, 2022) generally outperform traditional models through contextual embeddings and end-to-end learning, a simpler model was intentionally chosen to isolate the influence of the tokenization technique. The use of more complex architectures was avoided to reduce computational overhead and focus the evaluation on representational differences.

4.3 Lessons Learned

Key insights from the project include:

- **Preprocessing Matters:** Text normalization and cleaning improved model performance, even without deep learning.

- **N-grams Add Context:** Bi-grams and tri-grams helped capture basic contextual patterns important for sentiment.

- **BERT Tokenizer Helps:** Subword tokenization and pretrained knowledge enhanced performance, even without fine-tuning.

- **Model-Tokenizer Fit:** Best results likely require pairing advanced tokenizers with models that can fully utilize them.

5

In conclusion, although modern approaches provide superior accuracy, it has been demonstrated that with appropriate preprocessing and feature engineering, traditional pipelines can still deliver meaningful results.

References

Dipannita Biswas and K. M. Azharul Hasan. 2024. Adaptive bert-cnn model for efficient maintenance of variable length reviews in rating prediction. In *2024 27th International Conference on Computer and Information Technology (ICCIT)*, pages 1773–1778.

Archa Joshy and Sumod Sundar. 2022. Analyzing the performance of sentiment analysis using bert, distilbert, and roberta. In *2022 IEEE International Power and Renewable Energy Conference (IPRECON)*, pages 1–6.

Tomas Pranckevičius and Virginijus Marcinkevičius. 2016. Application of logistic regression with part-of-the-speech tagging for multi-class text classification. In *2016 IEEE 4th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, pages 1–5.

A   Appendix

[R1]   I read several of the reviews for this product, and in spite of some that indicated trouble, ordered. Here's my take on how you can improve your success rate with it. READ THE INSTRUCTIONS! Maybe they weren't included with the product reflected in some of the reviews, but they're in there now. They tell you twice not to fill the SlickDrip more than halfway and not to touch it, because...wait for it...it may collapse and you may make a mess and get burned. I didn't have any trouble with it. I'm willing to concede it may relax after while and if it does, I'll come back to eat crow.

Some reviewers indicate it drains too fast, producing weak coffee. Two factors may affect that, neither of which is in the instructions. However, the vendor sent an e-mail with detailed instructions on how to make the best coffee. Really nothing in there for you if you know how to make coffee. You need the correct grind, which is described as medium, "slightly larger than coarse sand." the correct grind is essential to the flow rate. The second factor is the filter. I haven't studied the porosity of filters, always using the same one. However, it can vary, and that will affect retention time. You may want to experiment. The more important consideration is to follow the instructions of the filter manufacturer. Did you ever read the box? Did you fold the filter on the crimped seams? This provides a strengthened rib along one edge that contacts the filter, and along the bottom seam. The weight of the water compresses it against the seam. Inasmuch as the water always has to flow through the doubled seam edge on the bottom, the flow rate is slowed.

I made great coffee with the SlickDrip. It's as good as anything from a Keurig, more similar to the results of the Chemex pour-over I own, and I'm a coffee snob. I'm very satisfied with it. I may buy another!

[R2]   Didn't purchase from Amazon but posting review. Ideally, this sounds perfect and is supposed to only increase you're electric about $10-$20 per month depending on how much laundry you do. It's just me so one load of laundry per week, except sheet week, so what can be wrong? This machine never dries!!!! First one I got took 12 hours to dry my clothes and they came out smelling burnt. I learned that apparently the first one was defective. The new one has now taken 6hours to dry. I understand it will take longer than regular machines but 6 hours is ridiculous. And I'm talking a regular one person load of small clothes and maybe a small towel! Six hours!!!!!! So I go to take the clothes out right away bc I notice they never smell clean and sure enough, they smell like dirty dish water, rather than the degergent and fabric softener I used. Oh, they are also a complete wrinkled mess, which I've never had an issue with on regular w/d units. Sorry, I hate ironing which is why I love things that can go in the dryer bc they don't come out wrinkled. I HATE THIS PIECE OF GARBAGE AND IT IS BEING RETURNED ASAP. Let's not talk about how much my electric bill jumped bc the machine was running for 12 hours last time. WTF? Great idea, but clearly something isn't right in the design.

[R3]   I'll be honest, I was very sceptical about a plastic washing machine. But this thing is just awesome! I have had it for 3 months now. I am a single male, and I use it about 2-3 times per week. I have had ZERO problems with it. I pulled it out of the box, put it in the shower surround, filled it with water and washed. There is NO assembly to do. And it is of quite good quality all around. It's not a bunch of cheap plastic that will fall apart in a year. This thing is so cool!

You can place this ANYWHERE there is a drain nearby. I live in a small studio apartment. I dont have a typical tub/shower. Just a shower surround. So

6

I just place it in the shower, drop the drain hose onto the floor, fill it with water and wash. Don't think it doesn't wash as well as a full size. It does! It is surprisingly quiet in both wash and spin cycles. Almost silent. So if you live in a multi-unit apartment complex, you dont need to worry about pissing off your neighbors.

It's a money saver too! Because in the wash cycle, the motor spins the tub in one direction for a few seconds then stops. It pauses for a few seconds, then spins in the opposite direction for a few seconds. So it use very little energy, whereas a full size washer uses a much bigger motor and it runs continuously for the entire wash cycle. There are 2 seperate motors in the unit. And the motors seem to be of good quality, and quite strong. I'm not worried about them failing after a year or so.
Another way it saves money is bc there is no heating element to dry clothes. BIG money saver right there! Instead, it has a seperate compartment which houses a stainless steel basket. The other motor spins the wet clothes so fast that it actually does a fantastic job of almost completely drying the clothes. It spins on a 0-5 minute automatic timer. More than enogh to do the job. (Note: full synthetic items come out almost completely dry. Only slightly damp to the touch. Cotton jeans and towels and the like, come out damp, and require some dry time. Depending on the climate where you live. If you live in a humid area, it may take a half day to dry the heavier cotton items)
Another way it saves money is that you really only add NO MORE than a tablespoon of laundry detergent. Seriously, that is PLENTY.

This 9.9 lb machine i bought, can do a pair of jeans and a few pairs of underware and socks in one load. Or one full size thick cotton bath towel and a few smaller items per load. I can easily fit full size bed sheets in it, but only top sheet, then bottem sheet in another load. But it handles it just fine. I'm pretty sure the 9.9 lb would handle a queen size. A full size comforter....No way, forget about it. Take that to the laundromat. (I think that one of the 13 lb machines will handle a comfortor though.) It will handle about 5-7 cotton t-shirts per load.

There is definetly more "hands on" work with one of these machines as compared to a full size automatic. Bc you do a wash cycle, (about 10-12 minutes) then you have to turn a knob to the "drain" setting, let it drain (takes about 20 seconds), then refill with water, turn knob to wash cycle, (which is really the rinse cycle. I do about 6 minute rinse), then come back and turn knob to drain. Then, move clothes over to the other compartment on the right, set timer to 5 minutes to spin clothes dry. As you are spinning, you can be filling up wash tub for next load. (YES, you can be washing and spinning at the same time bc compartments, settings, timers and motors are seperate systems). I know all that sounds like alot of work, but it really isn't. It's so quick and simple that it's almost fun! I bought some cheap extendable clothes hanging bars, that I put between the hallway walls to hang clothes on. When I'm done washing, I just pick up the unit and place it in the back hallway. The thing is super lightweight, that a 90 yo grandma can easily move it.

PROS:
-Cheap...$80.33 shipped to the house
-Good quality
-Very light weight
-Can be used in any dwelling where there is no waher/dryer hookups (perfect for apartments, studios, businesses, rv's/camp trailers etc)
-Uses almost no electricity. So it can be easily used with very small solar systems like you'd have in rv's etc
-uses very little detergent, no more than 1 tblsp per load
- Easily portable. So if you move, and the new place doesn't have W/D hookups....no problem
- Cleans every bit as good as a full size unit
- Very quiet operation in both wash and spin. Great for multi-unit apartment complexes.
- Excellent option if your full size breaks down and you cant afford to replace it right away
- Much smaller loads save water. Although you do more loads, I still feel like it would use less than a full size would use.
 - Very user friendly.
- Has a lint trap inside washer unit, that easily slides out/in for cleaning. And yes, it actually works surprisingly well. Traps quit a bit of lint.
- Solidly made
- Perfect for 1-2 people household

CONS:
- Smaller loads so its more hands on. Bigger size machines wouldn't have this con.

If you buy it and it turns out to be too small, exchange it, or give it to someone who needs it. Or donate it to a small business like an animal shelter, hair salon, food bank, charity organization etc.

Hope this helped!

[R4] What Amazon's product page says:

[Large-capacity space]This is a two-layer egg box, one layer can store 20 eggs, and a total of 40 eggs can be stored. The drawer design is dustproof and moisture-proof, so that each egg can be kept intact and fresh. Product Dimensions: 10.8 x 8.6 x 5.9 Inch.

[High-quality mate-

rial]This egg carton is made of high-quality food-grade plastic, PBA-free, thick and durable, and the transparent material allows you to clearly see the storage of the eggs.

[Space saving] The egg carton is compact and stackable, which can save a lot of space in the refrigerator, making the refrigerator look neat and orderly, and the groove design prevents it from being crushed or smashed.

 [Wide range of use] Egg carton is suitable for refrigerator, freezer, kitchen, dining table, storage room, camping and picnic. It can be easily put in the refrigerator and other places for easy portability.

[After-sales service]Our products are well tested, inspected and packaged before shipment. If you have any questions or suggestions about these egg tray, please feel free to contact us. Our service team will provide you with a 100% satisfactory answer within 24 hours.

What I say:

First off, my pet peeve: Why do all egg storage bins have such an odd number of eggs they hold? This one holds 40. If you purchase 3 dozen eggs, that is 36. So why 40? I just do not get it.

But that is a minor thing, really.

This 2-drawer bin is sturdier than most others that I have had in the past. It is well designed and the eggs within (unless they are small eggs) fit nicely so there is no rocking/rolling around in the bin to possibly break them - causing a MESS. Each egg is cradled in a dip. (So why would anyone purchase a bin where the eggs roll and break? It is just beyond my understanding.)

Easy to clean; however, I have filled mine up 3X now and it is still spotless. I did wash it once though, in a dish washing machine. I did not notice anything in the package saying that it was hand wash only and did not want to post that it was machine washable without testing it first. (If it does not hold up to multiple washing, I will come back and edit this post.)

It top portion is clear so you can easily see how many eggs you have left.

As usual, I will come back and edit this review if the product does not last long, is defective, or anything else.

[R5] [[VIDEOID:$4fd222938153f33...0720d$]] I'm so glad I bought this washer! I gave this product 5 stars because it really does the job. Clothes come out clean and nearly dry. I use less detergent. I don't have to drag my stuff 4-5 blocks. I don't have to worry about having enough money to do the wash (just enough to buy the washer n stuff, lol).

I'm grateful for the reviews and videos posted because they give great advise! CHECK THEM OUT; all of them. This product comes with some kind of instructions but no practical directions. The videos are filled with different types of practical advise. Some of the videos seem to be in not-so-small apartments that have washer hookups or standing drainage areas, lol. My apartment is small–750 sq ft–and has none of that. Don't be discouraged. This product is described as good for small apartments; and, it is. The dimensions are perfect. It's light enough to move back and forth and fit into small spaces. People who have difficulties lifting more than 20 lbs might have an issue if they need to lift it or carry multiple buckets of water. Not to worry, brainstorm.

Items you may need to purchase in addition to your washer: extension cord, inlet water hose, drainage hose, sturdy table, bucket or trash bin. Wait to receive your washer before purchasing any extra items.

* The washer is small and low; makes some but not a lot of noise. Do you have a sturdy table to put it on whether or not you'll be putting it in the bathtub? Do you have any ideas about where you want to put and use it?
* DO NOT leave your operating washer unattended! This machine will move–especially at the beginning and ending of your cycles; especially if loads are unbalanced! Is your table sturdy enough to deal with movement? Will your moving machine end up unplugging itself, moving other items, or falling off the table? How much time do you have to hold your machine in place while it does its job? Remain aware of what is going on with your machine.
* The electrical cord is short. Do you have an extension cord?
* The inlet hose doesn't fit all faucets, so have a plan. You can use a detachable shower head, kitchen faucet hose, or a bucket of water to fill your wash bin. You can purchase an appropriately-fitting hose; however, would you know what that is?? Why make yourself crazy?
If you have a detachable shower head, does putting the washer in the bathtub mean your electrical cord will cross over your faucet, shower head, or sit in the tub during draining cycle? Bad idea, unless you unhook the electrical cord each time and put it out of harms way OR are able to position washer and cord properly so water and cord don't overlap AND put the washer on a table so the electrical cord doesn't sit in your bathtub. Will you remember to do this EVERY TIME or at all? If you are anything like me,

that would be a NO. Don't risk damaging your machine before you finish your first load! Know yourself and do what is best for you considering all of your circumstances.
* The wash cycle has 5 timer settings which can be set to "soft" or "standard" agitation levels. It's up to you how long a wash cycle to run. Lifting the lid does not stop the wash cycle but does let you see whether the appropriate amount of detergent was used. You can add more, if needed. Standard agitation level is strong and works well. It might surprise you. Once wash cycle starts, you should NOT try to force or shorten the amount of time set. You will break your machine! Let it run; it will finish soon. Always start your timer at the least amount of time needed to wash whatever. You can always run it again after the cycle ends. You may be able to add time to your wash. I did this absentmindedly. Thank God, I didn't damage my machine. I think adding time to the washer cycle might be ok but I don't do that anymore. I just wash everything at 15 mins, lol.
* The drainage hose is too short, so have a plan: Will you be using the washer in the bathtub? Can the hose reach the sink while the washer is on the floor? If it does not, do you have a bucket or trash bin into which the dirty water can drain? Will you need to put the washer on a table?
I LOVE that the "drain" cycle (not the wash cycle) can be stopped at any time, so you can dump the water then return to draining. ** Don't forget to turn the knob each time to "standard" wash to stop the draining or you'll end up with water all over your floor–like me, lol.
* A wonderful tip I saw on one of the videos was about rinsing. After the draining cycle, turn knob to "standard" wash, refill the wash bin, run another wash cycle for maybe less time, then drain again. This is especially helpful if you put too much detergent and still have suds in your wash bin. Yup, it happens within the first few wash cycles as you learn how much detergent to use for the various types of materials being washed. Washing underwear usually requires less detergent than fuzzy pajamas or oversized towels.
Example #1: 10 underwear, one t-shirt, and a pair of shorts in a 15-min wash load took half of the wash bin requiring maybe a tablespoon of detergent. After washing then draining, refilled wash bin halfway, ran wash cycle for 9 mins then drained again. I divided the wash load in half; this load required 2 separate spin cycles of 5 mins each. Underwear were nearly dry; t-shirt and shorts, almost dry.

Time lapse: 34 mins plus machine set up.
Example #2: 2 oversized towels in a wash load took a little more than half the wash bin and required more detergent. Washed then drained, refilled wash bin a little more than halfway, ran wash cycle for 15 mins then drained again. Spun each towel separately requiring 2 separate spin cycles of 5 mins each. Towels were damp, no water dripping at all! Time lapse: 40 mins plus machine set up.
* The machine gets hot. So, you may want to fill your wash bin halfway to keep the motor or belt from burning out. This machine is not made with steel or other heavy metals–which is why it's so light. Therefore, overusing it or putting in very hot water may damage it. Be kind to yourself. Be kind to your machine. You have a washer in your place now. You can do another load later. Don't forget, you are standing there holding and / or watching your machine for however long. Don't waste your money or push the boundaries of your sanity by breaking your washer.
* The spin cycle is fabulous! Items come out very nearly dry, almost dry, or damp needing to be hung for a little while. Don't forget to put your machine in "drain" cycle to release the water from the spin cycle. The spin cycle also stops when you lift the lid which is helpful when balancing uneven loads. There are 2 lids, the outer and inner lid. Be careful. When closing the lids, close the inner lid first so that it doesn't snap off. Don't force it. When spin cycle starts up again, the machine will shake like crazy. Hold your machine!
* Between washes and draining, TILT your machine toward the drainage hose side and be ready with your bucket. It appears some water remains in the barrel within the machine. Drain it. I found out the hard way when I was moving my washer and found mysterious puddles of water on the floor. It happened a couple of times before I realized what was going on. I imagine that's why some customers complained of molding. If you don't drain it, the mold will come, lol. Wipe the inside of the lids and the wash and spin bins after use.
My experience: I was frustrated initially because needed a few things I didn't have: a table I could carry and use in either the kitchen or bathtub; a bucket or trash bin that wasn't too big so the water wouldn't be too heavy to lift every time I washed and drained; a mat to mute machine movement and sound so downstairs neighbor won't get frustrated; and drainage hose, inlet hose–neither fit my faucets. So, I bought those things. I wasted money on the

9

hose but everything else was great; I got it all on Amazon. I have a faucet hose in the kitchen sink. I put the washer on a table in the kitchen using the faucet hose to fill the wash bin. I use my trash bin, which doubles as a bucket, to drain the dirty water I dump into the sink. Machine setup and breakdown time is about 7-10 mins. Perfect! I highly recommend the SUPER DEAL Compact Mini Twin Tub Washing Machine with Spinner.