

Sub-category	No.	Feature Indicator	Description	Reference
Project	f1	<i>repo_age</i>	How long the project has existed on GitHub (in years)	[1]
	f2	<i>repo_has_org*</i>	Whether the type of owner of the project is an organization	
Maturity	f3	<i>repo_contributors*</i>	Number of contributors to the project	
	f4	<i>repo_subscribers*</i>	Number of subscribers to the project	
Project Activity	f5	<i>repo_num_open_issue*</i>	Number of open issues that existed at the time of issue creation for the project	
	f6	<i>repo_open_delta</i>	Time interval (in hours) between the issue creation and the closing date of the previous issue	[1]
	f7	<i>repo_closed_issue_time*</i>	Average length of time (in days) that the project has had issues closed at the time of issue creation	
	f8	<i>repo_num_commits_inproject</i>	Number of commits created in the project in the three months prior to issue creation	[2]
	f9	<i>repo_num_commits_project_T</i>	Number of commits in the project during the period of 2 weeks before the issue creation until the observation point T	[2]
	f10	<i>repo_num_issues_created_inproject</i>	Number of issues created in the project during the three months prior to issue creation	[2]
	f11	<i>repo_num_issues_created_inproject_closed</i>	Number of issues created and closed in the project in the three months prior to issue creation	[2]
	f12	<i>repo_num_issues_created_projected_T</i>	Number of issues created in the project during the period of 2 weeks before the issue creation until the observation point T	[2]
	f13	<i>repo_num_issues_created_project_closed_T</i>	Number of issues created and closed in the project during the period of 2 weeks before the issue creation until the observation point T	[2]
	f14	<i>repo_team_size</i>	Number of core developers in the project that closed issues during the three months prior to issue creation	[3]
Issue Content	f15	<i>issue_has_task_type*</i>	Whether the issue is considered to have a clear task type	
	f16	<i>issue_num_labels_T</i>	Number of labels added before the observation point T	[2]
	f17	<i>issue_text_similarity*</i>	Score of the Highest similarity of issues to previously closed issues within the project	
	f18	<i>issue_has_pr</i>	Is an issue a Pull Request	[4]
Issue Complexity	f19	<i>issue_cleaned_body_len</i>	Length of the combined title and body with markdown parsed and tags removed	[2]
	f20	<i>issue_num_codes</i>	Number of code snippets in the title and body of an issue	[5]
	f21	<i>issue_num_urls</i>	Number of URLs in the title and body of an issue	[5]
	f22	<i>issue_num_pics</i>	Number of images in the title and body of an issue	[5]
	f23	<i>issue_flesch_reading_ease</i>	Readability score (in Flesch Reading Ease) for the title and body of an issue	[5]
	f24	<i>issue_on_rest_day</i>	Whether the issue was created on the weekend	[1]
	f25	<i>issue_has_at_tag</i>	Whether the “@” tag exists in the title and body of an issue	[3]
	f26	<i>issue_has_conflict</i>	Whether the keyword “conflict” exists in the title and body of an issue	[3]
Issue Sentiment	f27	<i>issue_desc_positivity</i>	Positive sentiment score of the title and body of an issue	[4]
	f28	<i>issue_desc_negativity</i>	Negative sentiment score of the title and body of an issue	[4]
	f29	<i>issue_desc_pos_polarity</i>	Polarity sentiment score of the title and body of an issue	[4]
	f30	<i>issue_desc_subjectivity</i>	Subjective sentiment score of the title and body of an issue	[4]
Issue Activity	f31	<i>issue_hotness_T</i>	the hotness of an issue at the observation point T	[6]
	f32	<i>issue_num_comments_T</i>	Number of comments issue has received before the observation point T	[2]
	f33	<i>issue_mean_comment_size_T</i>	Average comment size of the comments received before the observation point T	[2]
	f34	<i>issue_has_core_comment_T</i>	Whether the question has comments from a core developer other than the creator before the observation point T	[3]
	f35	<i>issue_has_other_comment_T</i>	Whether the question has comments from non-core developer other than the creator before the observation point T	[3]
	f36	<i>issue_per_event_days_T*</i>	Average number of events per day for the issue at the observation point T	
	f37	<i>issue_num_assignments_T</i>	Number of assignment events before the observation point T	[2]
	f38	<i>issue_num_referenced_by_T</i>	Number of times the issue was mentioned in commit messages using the issue id, before the observation point T	[2]
	f39	<i>issue_num_subscribedby_T</i>	Number of persons subscribing to receive updates on the issue before the observation point T	[2]
	f40	<i>issue_num_actors_T</i>	Number of unique persons who have commented, referenced, or subscribed to the issue before the observation point T	[2]

Sub-category	No.	Feature Indicator	Description	Reference
Influence characteristics of creators	f41	<i>creator_age*</i>	How long the creator has existed on GitHub (in years)	
	f42	<i>creator_has_org*</i>	Whether the creator belongs to an organization	
	f43	<i>creator_followers*</i>	Number of followers of the creator	
	f44	<i>creator_newcomer*</i>	Whether the creator is a newcomer to the project	
	f45	<i>creator_core_member</i>	Whether the creator is a core developer of the project	[3]
Behavior characteristics of creators	f46	<i>creator_num_comments_T*</i>	Number of comments by the issue creator before the observation point T	
	f47	<i>creator_comments_positive_T</i>	Positive sentiment score of the comments of the creators before the observation point T	[3]
	f48	<i>creator_comments_negative_T</i>	Negative sentiment score of the comments of the creators before the observation point T	[3]
	f49	<i>creator_first_issue*</i>	Whether the issue is the first issue created by the creator within the project	
	f50	<i>creator_num_issues_inproject</i>	Number of issues created by the issue creator in the three months prior to issue opening	[2]
	f51	<i>creator_num_issues_closed_inproject</i>	Number of issues created by the issue creator that were closed in the three months prior to issue opening	[2]
	f52	<i>creator_num_issue_closed_inproject_T*</i>	Number of issues created by the issue creator that were closed during the period of 2 weeks before the issue creation until the observation point T	
	f53	<i>creator_issue_closed_days_T*</i>	Average length of closing time (in days) that issues created by the issue creator that were closed during the period of 2 weeks before the issue creation until the observation point T	
	f54	<i>creator_num_commits_inproject</i>	Number of total commits to the issue repository by the issue creator in the three months before the issue opening	[2]
Influence characteristics of participants	f55	<i>creator_num_commits_inproject_T*</i>	Number of total commits to the issue repository by the issue creator during the period of 2 weeks before the issue creation until the observation point T	
	f56	<i>participant_first_core_member*</i>	Whether the first participant is the core developer of the project	
	f57	<i>participant_num_T*</i>	Number of all participants in the issue before the observation point T	
	f58	<i>participant_association_score_T*</i>	Average association score of participants in the issue before the observation point T	
	f59	<i>participant_first_occured_T</i>	Whether the issue has a first participant at the observation point T	[3]
	f60	<i>participant_num_comments_T*</i>	Number of comments of all participants in the issue before the observation point T	
	f61	<i>participant_num_commits_T</i>	Total number of commits done by all participants to the project repository during the period from two weeks before the issue creation to observation point T.	[2]
	f62	<i>participant_num_unique_actors_by_commits_T</i>	Total number of unique participants who committed code to the project repository during the period from two weeks before the issue creation to observation point T.	[2]

NOTE: Features marked with _T are those whose values may change over observation time.

Features marked without * have been proposed in previous studies, while those with * are novel additions to the existing literature.

REFERENCES

- [1] D. A. N. d. Silva, D. M. Soares, and S. A. Gonalves, "Measuring Unique Changes: How do Distinct Changes Affect the Size and Lifetime of Pull Requests?," in Proceedings of the 14th Brazilian Symposium on Software Components, Architectures, and Reuse, 2020, pp. 121-130.
- [2] R. Kikas, M. Dumas, and D. Pfahl, "Using dynamic and contextual features to predict issue lifetime in github projects," in 2016 ieee/acm 13th working conference on mining software repositories (msr), 2016: IEEE, pp. 291-302.
- [3] X. Zhang, Y. Yu, G. Georgios, and A. Rastogi, "Pull request decisions explained: An empirical overview," IEEE Transactions on Software Engineering, 2022.
- [4] M. Izadi, K. Akbari, and A. Heydarnoori, "Predicting the objective and priority of issue reports in software repositories," Empirical Software Engineering, vol. 27, no. 2, pp. 1-37, 2022.
- [5] W. Xiao, H. He, W. Xu, X. Tan, J. Dong, and M. Zhou, "Recommending good first issues in GitHub OSS projects," in Proceedings of the 44th International Conference on Software Engineering, 2022, pp. 1830-1842.
- [6] A. B. Dhasade, A. S. M. Venigalla, and S. Chimalakonda, "Towards prioritizing github issues," in Proceedings of the 13th innovations in software engineering conference on formerly known as India software engineering conference, 2020, pp. 1-5.