List of features used

#	Features generated	Description
0	**AmountEUR**	Amount receivable in euros.
1	***days_late****	Number of days invoice was late, 0 if on time, negative if paid early, postitive if late.
2	***current_invoice_late_status****	Value indicating whether the invoice was late or not, where 1 means late, 0 means not late.
3	***days_between_created_and_paid****	Number of days between when the invoice was created and when the invoice was paid.
4	***days_between_created_and_due***	Number of days between when the invoice was created and the due date.
5	***total_number_invoices***	Number of invoices in total prior to the creation date of a new invoice of a customer.
6	***total_sum_invoices***	Total sum of invoice amount of the invoice owner prior to the creation date of a new invoice of a customer.
7	***average_days_late_and_early***	Average late and early days of all invoices both paid and outstanding prior to a new invoice for a customer.
8	***average_days_late_total***	Average days late of all invoices both paid and outstanding prior to a new invoice for a customer.
9	***average_days_early_total***	Average days early of all invoices both paid and outstanding prior to a new invoice for a customer.
10	***last_invoice_paid_status***	Value indicating whether the last invoice was paid or not; where 1 means paid, 0 means not paid, -1 means non-existent
11	***total_paid_invoices***	Number of paid invoices prior to the creation date of a new invoice of a customer.
12	***sum_amount_paid_invoices***	The sum of the base amount from all the paid invoices prior to a new invoice for a customer.
13	***total_invoices_late***	Number of invoices which were paid late prior to the creation date of a new invoice of a customer.
14	***total_invoices_early***	Number of invoices which were paid early prior to the creation date of a new invoice of a customer.
15	***sum_amount_late_invoices***	The sum of the base amount from all the paid invoices which were late prior to a new invoice for a customer.
16	***sum_amount_early_invoices***	The sum of the base amount from all the paid invoices which were early prior to a new invoice for a customer.
17	***total_outstanding_invoices***	Number of the outstanding invoices prior to the creation date of a new invoice of a customer.
18	***total_outstanding_late***	Number of the outstanding invoices which were late prior to the creation date of a new invoice of a customer.
19	***sum_total_outstanding***	The sum of the base amount from all the outstanding invoices prior to a new invoice for a customer.
20	***sum_late_outstanding***	The sum of the base amount from all the outstanding invoices which were late prior to a new invoice for a customer.
21	***average_days_late***	Average days late of all paid invoices that were paid late prior to a new invoice for a customer.
22	***average_days_early***	Average early days of all paid invoices that were paid early prior to a new invoice for a customer.
23	***average_days_outstanding_late***	Average days late of all outstanding invoices that were late prior to a new invoice for a customer
24	***std_days_late***	Standard deviation of late days for all invoices that were paid late prior to a new invoice for a customer.
25	***std_days_early***	Standard deviation of early days for all invoices that were paid early prior to a new invoice for a customer.
26	***std_days_outstanding_late***	Standard deviation of days late of all outstanding invoices that were late prior to a new invoice for a customer.
27	**ratio_1_late**	Ratio of 13 over 11. (total_invoice_late / total_paid_invoices). Ratio of paid invoices that were late
28	**ratio_1_early**	Ratio of 14 over 11. (total_invoices_early / total_paid_invoices) Ratio of paid invoices that were early
29	**ratio_2_late**	Ratio of 15. over 12. (sum_amount_late_invoices / sum_amount_paid_invoices). Ratio of sum of paid base amount that were late
30	**ratio_2_early**	Ratio of 16. over 12. (sum_amount_early_invoices / sum_amount_paid_invoices). Ratio of sum of paid base amount that were early
31	**ratio_3_late**	Ratio of 18 over 17. (total_outstanding_late /total_outstanding_invoices). Ratio of outstanding invoices that were late.
32	**ratio_4_late**	Ratio of 20. over 19. (sum_late_outstanding / sum_total_outstanding). Ratio of sum of outstanding base amount that were late
33	**client_cluster**	Cluster the client belongs to (Result Kmeans clustering on all the other features used in training).
33	**Lateness Category***	Arbitraty categories based on how many days the invoice was late

*** Not included in training**

****Most late features & descriptions are from other papers. Descriptions of features for early dates are modified accordingly. Sources:**

1. Appel, A. P., Oliveira, V., Lima, B., Malfatti, G. L., de Santana, V. F., & de Paula, R. (2019). Optimize cash collection: Use machine learning to predicting invoice payment. arXiv preprint arXiv:1912.10828.

2. Hu, P. (2015). Predicting and improving invoice-to-cash collection through machine learning (Doctoral dissertation, Massachusetts Institute of Technology).

3. Zeng, S., Melville, P., Lang, C. A., Boier-Martin, I., & Murphy, C. (2008, August). Using predictive analysis to improve invoice-to-cash collection. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1043-1050).