

### Feature Extraction

Feature Name	Procedure to Obtain Feature Value	Feature Value
Weekend or holiday expenses?	<b>Input:</b> extracted invoice date  <b>Output:</b> if the invoice date falls on a Weekend, return 1. Else, return 0.	0/1
Submitting the Same Invoice Multiple Times?	<b>Input:</b> extracted Invoice Date, Invoice Amount from User input, extracted Invoice ID, Reimbursement History data in database <b>Method:</b> Check if the current invoice is duplicated with any of the existing claims in the database. <b>Output:</b> If there are duplicated claims in reimbursement request history, return 1. Else, return 0.	0/1
Unusual spending – Whether the employee is in the claimed project?	<b>Input:</b> current employee ID, Project name from user input, ProjectEmployeeRelation table from database.  <b>Method:</b> Check if a person is associated with that project.  <b>Output:</b> Return 0 if yes. Else return 1.	0/1
Amount overclaim ?	<b>Input:</b> Invoice Amount Extracted, Invoice Amount from user input  <b>Method:</b> variance between invoice amount and typed amount  <b>Output:</b> Return 0 if no variance. Else, return 1.	0/1
Amount is multiple of 100?	<b>Input:</b> Invoice Amount typed  <b>Output:</b> Return 0 if no. Else, return 1.	0/1
Rounding amounts repeatedly claimed by the same person?	<b>Input:</b> Invoice Amount, Employee ID <b>Method:</b> check if the employee is repeatedly claiming rounding numbers in the most recent few claims in the reimbursement history. <b>Output:</b> Ratio of rounding amount claimed within the last three claims from the same employee.	0~1

likelihood of personal expense (Violations of Company Policies: Claims that violate company reimbursement policies, such as claims for personal expenses or expenses not related to work.):	<b>Input:</b> text of invoice <b>Method:</b> look for whether the text of invoice contains words <b>similar to</b> words such as "grocery", "supermarket", "tuition", "rent", "housing""Car", etc. <b>Output:</b> Return 0 if no, 1 if yes	0/1
Sudden Changes in Behavior: (Abrupt changes in reimbursement behavior, such as a sudden increase in claims)	<b>Input:</b> invoice amount from user input, future invoice date, employee ID, expense category, historical claim data from the ReimbursementClaimHistory table. <b>Method:</b> check if the current user input invoice amount has a big variance with predicted invoice amount for the current invoice date for this employee for the same expense category. <b>Output:</b> scaled variance between current invoice Amount and the predicted amount for same expense category for same employee	0~1
Invoice Date falls outside of Project Duration Dates	<b>Input:</b> Project Duration_startDate, ProjectDuration_endDate,current Invoice Date <b>Output:</b> Return 0 if invoice date falls in project duration.	0/1