Underground Metro Drivers: Occupational Problems and Job Satisfaction

Seyed Alireza Samerei · Kayvan Aghabayk · Mohamad Hosein Akbarzade

Abstract The performance of employees in a transport set is highly affected by their satisfaction and occupational problems. Underground metro drivers work in special conditions as compared with the drivers of suburban trains and urban tramways. Even though several studies have focused on the ergonomics of train drivers, there is a dearth of research on underground metro drivers. This study was conducted to investigate the factors affecting the underground metro drivers’ occupational problems such as depression and level of satisfaction. The study was carried out on underground metro drivers in Tehran, Iran. The required data were collected through interviews, self-report tests and a depression scale. The classification and regression tree (CART) model was used to measure the relationship among depression, different conditions and other demographic characteristics. By using hierarchical clustering, metro drivers were divided into two categories with desirable and undesirable levels of satisfaction with regard to varying levels of satisfaction in different fields. By labelling them and reusing the CART model, the most important parameters affecting the satisfaction of drivers were extracted. The results indicated that more than half of metro drivers were depressed. The most important factors in the incidence of depression were: resentment of some psychological factors such as dissatisfaction with loneliness in the cabin, stress and fear of an incident, fatigue, dissatisfaction with the resting place, work experience more than 3.5 years, dissatisfaction with rights and the need for a second job. Also, the most important factor affecting satisfaction was the quality of the resting place.

Keywords Underground metro driver · Occupational problems · Depression · Job satisfaction

1 Introduction

The system in the metro lines of Tehran and other major cities of Iran such as Mashhad, Tabriz and Isfahan is a kind of traditional guidance operation. In this operation, the guidance system relies on the metro driver. The metro driver plays a key role in driving the train as well as opening and closing the doors. Previous studies have shown that this job is a complex career. The metro driver operates in the context of surrounding environmental observations, signs of maintenance, automated cognitive process, memory constraints and dynamic decisions [1]. The train drivers not only direct the trains and the doors at the stations, but also play an important role in many other hidden tasks, such as predicting, observing, interpreting and responding to events around them [2].

Satisfaction and problems of metro drivers can be considered in four categories: (1) psychological needs and work strain; (2) psychological problems after an accident; (3) the quality of sleep and fatigue; and (4) status of the metro driver cabin.
Psychological needs and work strain of the metro driver

Conducting and controlling the metro is a psychological task that requires long-term work and intense concentration to pay attention to signals and stimuli [3, 4]. In addition, due to the special steadiness of this profession in the long run, it will be very difficult to maintain a conscious mind [5, 6]. Failure to pay attention to the psychological needs of train drivers and metro drivers will lead to mistakes, and an accident may occur [7, 8]. With regard to the recent advances in the train and metro industry in terms of higher performance and speed, the bulk of studies have been conducted to optimize the performance of the rail industry drivers [9, 10]. Such changes will have a significant impact on the psychological needs of metro drivers. To understand this, a conceptual model has been presented based on other important career models [11–15]. This conceptual model illustrates the relationship between psychological characteristics of work, volume and workload, and psychological needs. Psychological features of work include mental tasks, decision making and psychological needs [13]. These characteristics and psychological needs can create short-term stress that might be followed by interventions [12–16]. These reactions include mental and nervous pressure, fatigue, sleepiness and decentralization. As a result, daily retrieval for drivers and metro drivers should be considered. The recovery should be such that the metro driver has the same amount of energy for the day as before, otherwise short-term pressures will have long-term effects [13–16].

Psychological problems of the metro driver after an accident

Previous studies have shown that train drivers are distressed after a human and train crash, up to a year later [17]. A study was conducted to assess the psychological consequences of an accident on train drivers. In this regard, 202 drivers with an accident record had been compared to 186 other drivers. In three stages, i.e. 15 days, 3 months and 1 year after the incident and using the General Health Questionnaire–28 (GHQ-28) and a standardized diagnostic interview, a significant difference was observed in GHQ-28 score and stress disorder 15 days after the incident between drivers with an accident record and other drivers. However, there was no significant difference in the 3-month and 1-year periods [18]. Another study was conducted on the human-train collision to determine the amount and duration of remaining stress, anxiety, depression and quality of life after the accident. In addition, the relationship between post-incident stress predictors such as quality of life, distress and lack of meaning of life has been analyzed. Seventy-three drivers at the onset, 71 drivers at the outset, and 49 drivers after 4 weeks of rehabilitation were evaluated using a questionnaire, and emphasis was made on the effectiveness of post-accident rehabilitation [19]. Another study was conducted on the pattern of post-traumatic disorder and post-traumatic stress on metro drivers who have experienced an accident with a train with respect to personality and environmental characteristics. A questionnaire was used for data collection, of which 79% were filled in. Drivers with an accident experience had expressed more unpleasant physical and psychological symptoms. The environmental and personal conditions had no relation to the level of disorder and stress after the accident [20].

The quality of sleep and fatigue of the metro driver

Fatigue is one of the effective parameters in risk management that is related to work and non-working factors [21]. The consequences of fatigue in road and air transport are well documented. The lack of studies on the fatigue of train and metro drivers is apparent compared to car drivers and aircraft pilots [22, 23]. Train drivers do not have a routine sleep pattern due to shift work, which may lead to double fatigue. In some countries, such as the United States and Australia, biological models are used to identify and evaluate the fatigue of train operators to identify the specific shift that entails great fatigue [22]. It is worth noting that according to previous studies, drivers who are unmarried, hold a university degree and have limited driving experience exhibit a low perceived ability to control fatigue [24]. An extensive study was conducted to determine the effect of fatigue on the lack of attention to the red signal, centered on train drivers. Reasons, consequences and strategies to combat fatigue were determined [25]. Furthermore, a study has been carried out on train drivers in one of the Australian networks for sleep and performance. Considering the 8-hour shift and the rest in the special cabins, the drivers had a shorter sleep time and low-quality sleep than rest at their own homes. Hence, cumulative sleep apnea occurs, but drivers maintain their vigilance until the end of the work cycle [26]. Nighttime sleep of train drivers is usually short, and it is usually disordered due to overnight shifts, which in turn causes sleepiness and consequent human mistakes. In a study, the effect of rhythmic sleep regulation was investigated for preventing sleepiness. By studying 134 train drivers, it was concluded that sleeping rhythm decreases dramatically by adjusting the sleep rhythm [27].

The status of the metro driver cabin

In 2012, the Service Technique des Remontées Mécaniques et des Transports Guidés (STRMTG) issued a handbook on driver ergonomics concerned with the cabin. The goal was to determine design minima to meet the metro driver’s satisfaction and other metro driver requirements. Three categories of requirements were presented, as follows: (1) view, (2) position of controls and (3) windshield and side windows. In this study, two locations of the metro driver (center and out of center) were surveyed by collecting observations and questionnaires from seven tram networks. Observations indicated that some subway drivers complain
about seat quality and some other pedal access. Also, the reflection of the control screen on the windshield is sometimes annoying [28]. Extensive studies have focused on subway drivers and their operational space in the cabin, including feedback from the cabin space and ideal seat positioning. According to the reports, many metro drivers complained of frequent fatigue and general discomfort despite enough facilities to change the preferred position of the chair by the metro driver. The most important results of the study were the installation of an adjustable armchair for the metro driver and the need for proper coverage on dashboards to reduce the reflection of light on the windshield of the cabin [29]. Meanwhile, the comfort level of the driver’s perception of the signs and signals in the cabin and the route affect safety [30].

The main objective of this study is to assess the satisfaction of the underground metro drivers with various aspects of physical and psychological aspects, and to examine the possibility of depression among them and the factors affecting it. Despite studies on the psychological aspects of train drivers in normal conditions or after an accident, limited studies have been conducted on underground metro drivers. Underground metro drivers work in special conditions as compared with the drivers of suburban trains and urban tramways. These special working conditions for underground metro drivers are as follows: continuous work in tunnels and cabins, stopping at nearby and overcrowded stations, safety concerns, especially in the stations, and other responsibilities. Moreover, metro driver faces many challenges. Some of the problems of metro drivers are loneliness in the cabin, limited space and light of the cabin and tunnel, long shift, uniformity of work, inadequate income, etc. Accordingly, due to the sensitivity of this occupation, examination of the dissatisfaction and probability of depression among metro drivers and the factors affecting them seem quintessential.

2 Methodology

The metro organization of Tehran was selected to study the conditions of metro drivers. Group meetings and initial single interviews were held with a number of drivers from Tehran Metro Company. In this way, their main problems and concerns were identified. According to the metro drivers, the main problems were noise pollution, loneliness, accidental stress, air conditioning and road lighting. Figure 1 shows the percentage of these factors. Parameters affecting the mental and physical aspects of the metro driver were extracted using preliminary interviews and previous studies. A questionnaire was prepared based on these parameters. Also, given the oral explanations and metro driver conditions, the probability of depression among them was studied. Depression is one of the most dangerous and important issues in the field of mental health of the community. Lack of attention to depression causes irreparable problems. For this purpose, a standard depression test was used, and the relationship between all psychological and physical parameters of the metro driver with depression was studied. Also, by categorizing metro drivers at two levels of satisfaction, effective parameters were determined for the satisfaction level of metro drivers.

2.1 Participants

Tehran Metro consists of five urban lines and one active suburban line with a total length of 204 km (127 miles). According to the data obtained in 2015, an average of 3 million passengers per day and 670 million passengers per year would travel on 1000 active wagons.

Data were collected through interview with Tehran’s metro drivers who were randomly selected. The sample of study consists of 96 Tehran metro drivers. Due to the fact that there is no female metro driver in Tehran, 96 male metro drivers have been studied. The average age of the metro drivers was 32.96 years with a standard deviation of 3.78 years, and the average of work experience (background) was 7.97 years, with a standard deviation of 2.92 years.

2.2 Self-Developed Questionnaire

The questionnaire was designed based on the previous studies on psychological requirements, the previous studies of hardware and environmental ergonomics of metro and train drivers, and also the major issues raised in the initial interviews with the metro drivers.

The items related to psychological requirements that have been measured in this questionnaire and reported in various studies were fatigue, dissatisfaction with loneliness

![Diagram](image-url)
in the cabin, stress from an accident occurrence, the post-
traumatic stress, occupational safety, need for a second job
and the disruption of family relationships.

The related hardware and environmental ergonomics
items are sport activity, the comfort of the metro driver’s
seat, rest time, cabin and route lighting, visibility, air
conditioning, cabin temperature and noise pollution.

Based on the abovementioned parameters, a question-
naire was prepared including multiple-choice questions to
measure the drivers’ occupational problems that deal with
the environmental, mental, physical and other social and
individual factors (first part of questionnaire) and degree of
satisfaction or dissatisfaction (second part of question-
naire). The respondents rated each item on a six-point
Likert scale (1=never, 2=very low, 3=low, 4=medium, 5
=high, 6=too high), and the difference between every two
successive degrees is the same. The internal consistency of
the scales as measured by Cronbach’s alpha was 0.82 for
the first part and 0.89 for the second part of the question-
naire, and the questionnaires were validated by a panel of
experts in the field of psychology.

2.3 Beck Questionnaire

To measure depression, the Beck Depression Inventory
prepared by Aaron T. Beck was used [31]. The Beck
Depression Inventory is a 21-item multiple-choice self-re-
port inventory among psychometric tests widely used to
measure the depression degree in adolescents and adults.
This questionnaire assigns 1–6 points based on the answers
given by people in response to the questions (1: non-de-
pression, 2: slightly depressed, 3: needing to refer to the
psychologist, 4: relatively depressed 5: severe depression, 6:
excessive depression). Points 1 and 2 are in the non-de-
pressed group, and from grade 3 upward are in the depressed
group, and they require a referral to a psychologist. Based on
the obtained information, more than half of the statistical
population of the study was suffering from depression. Fig-
ure 2 shows the ratio of the depressed and non-depressed of
the statistical population. The internal consistency of the
scales as measured by Cronbach’s alpha was 0.91.

3 Models and Data

3.1 Classification and Regression Tree

Parameters affecting the possibility of depression of metro
drivers were extracted by dividing them into two groups:
depressed and non-depressed. For this purpose, a classifi-
cation and regression tree algorithm (CART) model was
used to examine the relationship between independent
parameters examined through a questionnaire with
dependent depression parameters. Figure 3 depicts the basis
of the CART model. The principle of the CART method in
developing the classification tree is such that, at first, the
entire data is concentrated in the upper node of the tree,
which is called the root node. The data are divided by one of
the independent variables (splitter) that produces the most
purity. In fact, the data in each child node is more homoge-
nous than its upper parent node. This process continues to the
extent that the data in each node has the highest purity. It is
referred to as the terminal node or the leaf.

The most popular criterion for node purity is the Gini
index, which is obtained from the following equation:

\[
Gini(t) = \sum_{i \neq j} \frac{p(j|t)p(i|t)}{p(t)}
\]

where \(i\) and \(j\) are categories of the target field, and
\[
p(j|t) = \frac{p(j,t)}{p(t)}, p(j,t) = \frac{\pi(j)N_j(t)}{N_t}, p(t) = \sum_j p(j,t)
\]

where \(\pi(j)\) is the prior probability value for category \(j\),
\(N_j(t)\) is the number of records in category \(j\) of node \(t\) and \(N_t\)
is the number of records in category \(j\) of the root node.
Note that when the Gini index is used to find the
improvement for a split during tree growth, only those
records in node \(t\) and the root node with valid values for the
split predictor are used to compute \(N_j(t)\) and \(N_t\), respec-
tively. The Gini index is equal to 0 when all the
observations in one node belong to a unique group, which
shows the least impurity.

3.2 Hierarchical Clustering

A hierarchical clustering (HC) algorithm had been used in
order to cluster the metro drivers in terms of satisfaction in
different fields. HC has two types of strategies, divisive and
agglomerative. Divisive methods are “top-down”
approaches in which all records start in one cluster and splits are performed recursively as one moves down the hierarchy. Agglomerative methods are “bottom-up” approaches in which each record starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. The Ward linkage algorithm, which is one of the agglomerative methods and has widely been used in many studies [32], was employed in the current study to group the metro drivers. Figure 4 represents the basis of the hierarchical clustering model. In this figure (dendrogram), the number of clusters obtained by defining a threshold can be seen.

### 3.3 Variable Significance

The identification of variables has major importance in the prediction of the target variable and is one of the most crucial steps in modeling. The importance of variable $x_i$ that intervenes in the CART model is defined in the following equation:

$$VIM(x_i) = \sum_{t=1}^{T} \frac{n_t}{N} \Delta Gini(S(x_i, t))$$

where $\Delta Gini(S(x_i, t))$ represents the reduction of the Gini index on the basis of variable $x_i$ in node $t$, $n_t/N$ is the proportion of the observations in the dataset that belong to node $t$, $T$ is the total number of nodes and $N$ is the number of total cases. This value is calculated for all of the independent variables and is scaled such that its summation for all variables is one. The variable that has the most importance with respect to the others has the largest number [33, 34].

### 3.4 Rules Extraction

The decision tree’s structure can be transformed into rules in order to extract potentially useful information. Each possible path of the decision tree, from root to leaf, can be a rule, and its validity can be measured by support and lift criteria. The support is the ratio between the number of items belonging to the class $j$ of the terminal node $t$ and the total number of items. The lift is the ratio between the proportion of items in the terminal node $t$ and root node that belong to the class $j$.

### 3.5 Model Assessment

In this study, accuracy is measured by tenfold cross-validation for CART models, due to the limited number of data. Also, overall precision, recall and $F_1$ score are extracted according to the classification table to classify the dependent variable in two levels (e.g., positive and negative). Accordingly, four different cases are summarized in Table 1.

As illustrated in Table 1, TP and TN denote the number of positive and negative samples, which were accurately predicted by the model. In return, FP and FN indicate the number of positive and negative samples, which were wrongly classified. To assess the model performance using
the classification table, the following measurements should first be calculated:

\[
\text{Accuracy} = \frac{TP + TN}{Pos + Neg} \tag{4}
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{5}
\]

\[
\text{Sensitivity or Recall} = \frac{TP}{TP + FN} \tag{6}
\]

\[
F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}
\]

Equation (4) can be used to obtain the model’s overall accuracy. The closer the model is to 1 in accuracy, the stronger it is in classifying the samples. Equation (5) can be used to obtain the model’s precision, and it is the ratio of correctly predicted positive observations to the total predicted positive observations, which shows the accuracy of the model. Also, high precision relates to the low false positive rate. Equation (6) can be used to obtain the model’s recall, and it is the ratio of correctly predicted positive observations to all observations in the actual class, which shows the sensitivity of the model. The \( F_1 \) score might be a better measure to use if we need to seek a balance between precision and recall, especially to choose the best decision tree depth. Equation (7) can be used to obtain the model’s \( F_1 \) score.

### 3.6 Data

By collecting questionnaires completed by metro drivers, a database of 96 rows of information was prepared. Each of the information rows contained 30 independent parameters, and Table 2 shows descriptions of each variable. Eventually, 30 independent variables and one target variable (depression or non-depression) were inserted into the CART model. Also, 17 parameters, which were related to the satisfaction level of metro drivers, were inserted into the hierarchical clustering model. In summary, in order to identify the parameters affecting the overall satisfaction of subway drivers, only 17 independent parameters related to satisfaction in different fields were entered into the model, and the parameter of satisfaction with rest place was significantly different from the others. It has affected the overall satisfaction of the metro drivers. For the depression model, 30 independent parameters were included in the model, 17 of which were for partial satisfaction, and 13 were for psychological and demographic ones. Therefore, partial satisfaction is shared between the two models of satisfaction and depression as model input data to observe the independence among the 17 independent parameter variables.

### 4 Results

#### 4.1 Depression Model

The CART method was used to implement the decision tree, analysis and classification of metro drivers into two categories of depressed and non-depressed. In modelling, we usually face the over-fitting problem so that the model has a very high precision but does not work well in the evaluation process. One way to prevent over-fitting in decision trees is to decide whether or not to prune, control and prevent the increase in the depth of the tree. One of the parameters is the \( F_1 \) score controller, which, if maximized, represents the balance between precision and recall of the model and overcomes the problem of over-fitting to a great extent. With multiple modelling with different depths, accuracy and \( F_1 \) score, the three-step expansion depth of the decision tree was selected. The decision tree made is displayed in Fig. 5. In this model, the variable that creates the most purity in the data separation, or in other words, the most reduction of the Gini index, is located at the highest level of the tree. Also, the next variables are determined based on the most reduction in the Gini index in the other nodes of the decision tree. Meanwhile, variables that are at higher levels of the decision tree are more important.

Given the limited numbers of data, using a tenfold cross-validation method, prediction models of depression probability were developed ten times and evaluated. The mean accuracy of the model was 0.72, which is appropriate. In Tables 3 and 4, the evaluation of each step and the final confusion matrix of all the steps are displayed. The accuracy of diagnosis of depression and lack of depression in this model was 61% and 79%, respectively, and the model’s evaluation shows that the input parameters of the model were capable of detecting depressed and non-depressed metro drivers.

#### 4.1.1 Structure of the Depression Model

A decision tree consists of multiple paths from root to leaf representing a classification decision rule. In this section, we tried to explain the nodes, the percentage of data classification and purification obtained in each node, and also considered all paths of the decision tree. In order to
better understand the relationship between the text and the shape of the decision tree, the paths leading to each node and the leaf enclosed in parentheses are summarized. The tree is made up of 13 nodes and 7 leaves (Fig. 5). Root nodes are divided into two groups based on the variance parameter of loneliness. Drivers with resentment of loneliness higher than 4 (5=high, 6=too high) were more likely to be depressed than those with a resentment rate of less than 4 [85% (29/34) vs. 37% (23/62)]. Among the metro drivers with a high level of resentment (node 1), fatigue is a parameter that influences the differentiation of the depressed and non-depressed drivers. The metro driver with fatigue rate more than 2, had 87% chance of depression (node 0, 1, 3), and otherwise 100% non-depressive (node 0, 1, 4). In node 3, the effective factor is the work experience of the metro driver. The interesting thing

Table 2  Variable description

| Variable | Description | Used in CART | Used in HC | Questionnaire part |
|----------|-------------|--------------|------------|--------------------|
| Fatigue rate | Ordinal<sup>a</sup> | ● | | 1 |
| Education level | Categorical<sup>a</sup> | ● | | 1 |
| Age | Continuous | ● | | 1 |
| Background | Continuous | ● | | 1 |
| Sport activity | Ordinal<sup>a</sup> | ● | | 1 |
| Trip | Ordinal<sup>a</sup> | ● | | 1 |
| Crash experience | Categorical<sup>b</sup> | ● | | 1 |
| Post-traumatic stress | Ordinal<sup>a</sup> | ● | | 1 |
| Disruption of family relationships | Ordinal<sup>a</sup> | ● | | 1 |
| Second job | Categorical<sup>b</sup> | ● | | 1 |
| Second job requirement | Ordinal<sup>a</sup> | ● | | 1 |
| Problems | Categorical<sup>a</sup> | ● | | 1 |
| Satisfaction of the metro driver’s seat | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of cabin lighting | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of control room connection | Ordinal<sup>a</sup> | ● | ● | 2 |
| Dissatisfaction of loneliness in the cabin | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of route lighting | Ordinal<sup>a</sup> | ● | ● | 2 |
| Stress rate from an accident occurrence | Ordinal<sup>a</sup> | ● | | 2 |
| Satisfaction of occupational safety | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of visibility | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of air conditioning | Ordinal<sup>a</sup> | ● | ● | 2 |
| Dissatisfaction of noise pollution | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of rest room | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of contact with the traveler | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of rights | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of cabin temperature | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of rest time | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of time schedule | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of applying abilities | Ordinal<sup>a</sup> | ● | ● | 2 |
| Satisfaction of relationship with other drivers | Ordinal<sup>a</sup> | ● | ● | 2 |
| Depression | Categorical<sup>e</sup> | ● | - | |
| Cluster | Categorical<sup>f</sup> | - | | |

<sup>a</sup> Ordinal: 1=never, 2=very low, 3=low, 4=medium, 5=high, 6=too high
<sup>b</sup> Categorical: 1=yes, 2=no
<sup>c</sup> Categorical: 1=noise pollution, 2=loneliness, 3=stress of accident, 4=air conditioning, 5=route lighting
<sup>d</sup> Categorical: 1=diploma, 2=bachelor of science, 3=master of science, 4=PhD
<sup>e</sup> Categorical: 1=non-depressed, 2=depressed
<sup>f</sup> Categorical: 1=cluster 1, 2=cluster 2
is the threshold less than 3.5 years at this node. Drivers with work experience of more than 3.5 years, which is considered to be a short time, were more likely to have depression with probability of 90% (node 0, 1, 3, 7).

In node 2, we see a very high threshold for data segregation, which means that there is a possibility of depression only in extreme stress situations, with a probability of 67%. If there is relative satisfaction in terms of loneliness and lack of stress, satisfaction of the resting place has been effective in distinguishing the possibility of depression (node 6). The metro drivers were not depressed, with a probability of 87%, if they had a level of satisfaction of more than 1 with the resting place. At node 5, if the rate of satisfaction with salary and benefits is greater than 2, the likelihood of non-depression is 58%; otherwise the metro driver will be depressed with the probability of 100%.

4.2 Satisfaction Model

The clustering method was used to study the factors affecting satisfaction and classification of metro drivers. In this way, metro drivers with the same level of satisfaction will be within one cluster and will be separated from others. By defining a threshold in the dendrogram, one can define the desired number of clusters. Regarding Fig. 6, two general clusters are defined under the rubric of desirable satisfaction (cluster_1) and undesirable satisfaction (cluster_2). Accordingly, each metro driver has a label either cluster_1 or cluster_2.

To find the effective factors on this clustering, decision tree modelling was used. In the decision tree model, satisfaction variables are defined as inputs of the model, and cluster labels as dependent variables. The general shape of the decision tree is shown in Fig. 7. In this model, the variable that creates the most purity in the data separation, or in other words, the most reduction of the Gini index, is located at the highest level of the tree. Also, the next
variables are determined based on the most reduction in the Gini index in the other nodes of the decision tree. Meanwhile, variables that are at higher levels of the decision tree are more important.

Due to the limited numbers of data, a tenfold cross-validation method, a satisfaction predictive model, was developed and evaluated ten times. The mean accuracy of the model was 0.86, which is evaluated as very appropriate. In Tables 5 and 6, the evaluation of each step and the final confusion matrix of all the steps are displayed. The accuracy of satisfaction and dissatisfaction in this model was 84% and 88%, respectively. The evaluation of the model shows that 17 input parameters of the model had the ability to predict cluster_1 and cluster_2.

4.2.1 Structure of the Satisfaction Model

A decision tree consists of multiple paths from root to leaf representing a classification decision rule. In this section, we tried to explain the nodes, the percentage of data classification and purification obtained in each node, and also considered all paths of the decision tree. In order to better understand the relationship between the text and the shape of the decision tree, the paths leading to each node and the leaf enclosed in parentheses are summarized.

The decision tree was used to identify the effective parameters of driver segregation in two desirable and undesirable satisfaction groups (Fig. 7). Satisfaction with resting place was very effective in their initial segregation (node 0). If the driver’s satisfaction rate from the resting place is greater than 2, it is likely to have a desirable satisfaction cluster (95% vs. 14%).

For metro drivers with a satisfaction rate of more than 2 (node 1), if the satisfaction rate of communication with the control room is more than 2, with the probability of 97.5%, the metro driver might be in the satisfaction group (node 0, 1, 3). In node_3, the determinant is air conditioning. If the satisfaction with the cabin air conditioner is more than 3, the metro driver might be satisfied with the probability of 100% (node 0, 1, 3, 7).

On the other hand, in case of dissatisfaction with the resting place (node 2), the cabin air conditioning is a separating parameter. If the satisfaction rate for cabin air...
conditioners is less than 3, then the metro driver might be satisfied with the probability of 94% (node 0, 2, 6). At node_5, the satisfaction rate of more than 1 in the field of path lighting may keep the metro driver satisfied (node 0, 2, 5, 10). In node_6, the only factor that could keep drivers satisfied is their salaries and benefits. If the rate of satisfaction with rights is greater than 1, the metro driver might be in the satisfaction group (node 0, 2, 6, 11); otherwise, they might be dissatisfied with the probability of 96%.

### 4.3 Variable Importance

#### 4.3.1 Depression

One of the advantages of CART models is the ability to detect important and effective parameters in predicting the model input labels. The importance of each parameter depends on the ability of that parameter to purify the data. The significance of the important parameters of the predicted model of depression of metro drivers, which had 30 inputs, was estimated using Eq. 3. The importance of each data was normalized, sorted and is given in Table 7 in order of importance. Parameters with little importance were removed from the table.

The lack of satisfaction of metro drivers with loneliness in the cabin was the first and most important parameter affecting the probability of depression of metro drivers with 100% significance. Metro drivers spend most of their time in the cabin alone. The relationship between loneliness and depression has been studied in most studies in other areas of psychology, and the correlation between the two has been confirmed [35, 36]. In this area, we also observed the relationship between loneliness and depression.

Stress and fear of an incident was the second most important parameter affecting depression, with 74.8% degree of importance. In the case of metro, despite the relative isolation of the route, except at the stations, the fear of collision has been detected as influential for metro drivers. Extensive research has been conducted on train drivers with regard to the stress of an incident or the disasters caused by past events. This issue has great importance due to the nature of railway interconnections with humans along the way and has a significant impact on the performance of metro drivers. Metro drivers need special vigilance to prevent accidents. Physical and psychological effects have been reported after an accident for train drivers, who emphasize the need to improve the

| Table 5 Tenfold cross-validation results of the satisfaction model |
|---------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tenfold 1 2 3 4 5 6 7 8 9 10 Mean value |
| Accuracy 0.9 1.0 0.8 0.9 0.8 0.8 0.8 0.77 0.88 0.77 0.86 |

| Table 6 Confusion matrix and accuracy of the satisfaction model |
|---------------------|-----|-----|-----|-----|-----|-----|
|                      | Positive | Negative | Accuracy (%) | Precision | Recall | F1 score |
| Positive 43 8 84 0.89 0.84 0.86 |
| Negative 5 40 88 |
| Overall percentage 50% 50% 86 |

| Table 7 Variable significance of the depression model |
|-----------------------------|-----------------------------|-----------------------------|
| Variables                   | Normalized importance (%)   | Variables                   | Normalized importance (%)   |
| Dissatisfaction of loneliness in the cabin | 100 | Satisfaction of control room connection | 16.8 |
| Stress rate from an accident occurrence | 74.8 | Posttraumatic stress | 16.6 |
| Fatigue | 50 | Satisfaction of relationship with other drivers | 14.7 |
| Satisfaction of resting place | 47.65 | Satisfaction of the metro driver’s seat | 13 |
| Background. | 46.3 | Dissatisfaction of noise pollution | 9.8 |
| Satisfaction of rights | 31 | Satisfaction of applying abilities | 7.94 |
| Second job requirement | 18 | Disruption of family relationships | 4.2 |
mental conditions of train drivers following an accident [18–20] and emphasize the rapid treatment of metro drivers involved in an incident by a psychologist in order to return to work [37].

The fatigue and satisfaction with the resting place with 50% and 47.65% degree of importance, respectively, were the third and fourth effective parameters. These criteria were likely to be affected by the resting place of the metro drivers, the amount of rest, the amount and the quality of sleep of drivers during the day. It is worth noting that, according to previous studies, these factors did not have much effect on performance [26], but according to our findings, they had a great effect on depression. The number of shifts, work strains and metro driver responsibilities affected the feeling of tiredness [25]. Sleepiness was one of the criteria for fatigue. Sleep rhythm modification prevents metro drivers from sleeping [27].

The background of metro drivers was one of the significant parameters influencing the predictive model of depression. According to our findings, metro drivers might be depressed after 3.5 years of work. This is a very alarming issue. The monotony of the work of these people during the day and the movement of a constant and recurring route might increase the likelihood of depression of metro drivers. Monotony is an inherent characteristic of rail industries, and a relatively minor increase in cognitive demand can mitigate adverse monotony-related effects on performance for extended periods of time [38].

The parameters of satisfaction with the received salary and need for a second job in metro drivers were influential on the likelihood of depression. This parameter was lower in some studies. This is probably due to the conditions of society and the current economic problems, which puts much psychological pressure on most people, including metro drivers. Economic problems can lead to dissatisfaction with working conditions and the need for a second job. If metro drivers have different jobs, the amount of rest and sleep required for their recovery will be affected.

Other important parameters with less importance included satisfaction with the quality of communication with the control room, posttraumatic stress, satisfaction with the level of communication with other metro drivers, the seat of the driver, the annoyance of noise pollution, the satisfaction of the application metro drivers’ abilities and dissatisfaction with family disruption. These parameters will be addressed in the next priority if needed.

4.3.2 Satisfaction

The significance of the important parameters of the model for predicting the level of metro drivers’ satisfaction was estimated using Eq. 3. The importance of each data was normalized and sorted, and is shown in Table 8 in order of importance. Based on the results, the quality of rest areas was the most important and effective parameter in determining the satisfaction level of metro drivers. The notable point in this regard is the lack of expression of this issue as the major problems of metro drivers in self-reporting questionnaires. The quality of resting places was considered as a hidden and very influential factor on the level of satisfaction. Despite the emphasis on the high impact of income and rights on job satisfaction in post-transition economies like Iran [39], this parameter was not very important for metro drivers. The other effective parameters with the lower degree of importance included satisfaction with rights, air conditioning, route lighting, quality control and cabin lighting. These parameters will be addressed in the next priority if needed.

4.4 Rules Extraction

Each of the paths leading to the tree leaves can be a rule. Measuring the ability to rely on these rules is important. In this study, these standards were measured using the support and lift criteria. In this study, we set the support and lift threshold to 30% and 1.5, respectively. The most important rule that most likely causes depression to be taken seriously was dissatisfaction with loneliness at rates higher than 4 and fatigue at a rate higher than 2 and a background of more than 3.5 years as a metro driver. Valid rules for predicting the likelihood of a metro driver’s depression or lack of depression are given in Table 9.

The most authoritative rule that most likely results from the metro driver’s dissatisfaction is the satisfaction with the resting place, air conditioning and benefits, at rates lower than 3, 3 and 2, respectively. Valid rules for predicting the level of satisfaction of the driver are given in Table 10.

| Variable                  | Normalized importance (%) |
|---------------------------|---------------------------|
| Satisfaction of resting place | 100                       |
| Satisfaction of rights    | 9.5                       |
| Satisfaction of air conditioning | 6.5                      |

| Variable                  | Normalized importance (%) |
|---------------------------|---------------------------|
| Satisfaction of route lighting | 6                         |
| Satisfaction of control room connection | 5                        |
| Satisfaction of cabin lighting | 4.8                      |
Table 9 Most important rules of the depression model obtained based on the tree

| Rules: IF… | Then | Support | Lift |
|------------|------|---------|------|
| IF (dissatisfaction rate of loneliness in the cabin [background > 3.5 years] > 4) AND (fatigue rate > 2) AND (stress rate from an accident occurrence ≤ 5) AND (satisfaction rate of resting place > 1) | Depressed | 29/96 × 100% = 30.2% | 29/96 / 32/96 = 1.67 |
| IF (dissatisfaction rate of loneliness in the cabin ≤ 4) AND (stress rate from an accident occurrence ≤ 5) AND (satisfaction rate of resting place > 1) | Non-depressed | 44/96 × 100% = 30.2% | 44/96 / 48/96 = 1.91 |

Table 10 Most important rules of the satisfaction model obtained based on the tree

| Rules: IF… | Then | Support | Lift |
|------------|------|---------|------|
| IF (satisfaction of rest room > 2) AND (satisfaction of control room connection > 2) AND (satisfaction of air conditioning > 3) | Satisfaction | 39/96 × 100% = 40.0% | 39/96 / 51/96 = 1.88 |
| IF (satisfaction of resting place ≤ 2) AND (satisfaction of air conditioning ≤ 2) AND (satisfaction of rights ≤ 1) | Dissatisfaction | 31/96 × 100% = 32.3% | 31/96 / 45/96 = 2.06 |

5 Discussion and Conclusion

Regarding the effect of mental health of the staff of a transportation complex on the function and level of safety, the prompt diagnosis of a mental illness is very important. In this study, according to modelling of depression, we found that the risk of depression significantly increased among metro drivers who experienced some mental health problems, including loneliness in the metro cabin, stress rate from an accident occurrence, fatigue, over 3.5 years of work experience and necessity to take a second job because of the inadequacy of their salaries and wages. Due to the conditions of underground subways, drivers are alone in the cabin for long hours in underground tunnels. These conditions are the 'mainstay' of subway drivers’ differentiation with other train and tram operators, which have been less addressed; however, in other social studies, the association between depression and loneliness has been confirmed [35, 36]. It should be noted that this study found that the most important influencing parameter in depression among the subway drivers was the extent of their suffering in the cabin alone. Stress and resentment from past events are common to subway and train drivers; such drivers are at high risk of developing panic disorder that may affect other physical and psychological conditions [20, 37, 40].

According to the results of this study, while confirming the existence of these cases among subway drivers, these factors will in turn lead to depression. The shifts of train drivers, which are sometimes uneven and irregular in their sleep rhythms and fatigue, affect performance and safety according to previous studies [21–25]. The fatigue parameter reported by metro drivers, while it may have functional problems, is also associated with the probability of depression.

Previous studies have identified some risk factors, including loneliness, stress and resentful feelings inflicted by past accidents and fatigue. In previous studies, little attention has been paid to work experience, which is highly controversial and indicates that metro drivers were exposed to risk factors for depression after work experience of 3.5 years. Also, metro drivers were exposed to risk factors for depression due to the economic conditions governing society, the inadequacy of drivers’ salaries, wages and their need for taking a second job.

To address some psychological factors associated with the experience of depression, we thus recommend the modified shift work, improving the resting places provided for metro drivers, taking into account periodic leaves, operating in different subway routes and stopping doing monotonous tasks, reforming and improving their salaries and wages and taking measures in metro cabins to reduce stress caused by loneliness. Furthermore, the stress caused by a rail crash and resentful feelings inflicted by past accidents may also increase the risk of depression among metro drivers. Therefore, counseling and psychotherapy sessions are needed for metro drivers who have experienced such accidents. Holding effective meetings and providing periodic entertainment for metro drivers with their families can eliminate problems and dissatisfaction related to the lack of proper relation between metro drivers and family disruption.

Also, the risk of depression significantly increased among metro drivers who were dissatisfied with their job due to the quality of their resting place, the quality of communication with the metro control room, the metro driver’s seat location and the noise pollution that can be easily solved by the standards related to wagons and infrastructures. In the case of train drivers, the complexity...
and multiplicity of tasks have resulted in reduced performance on one hand and the likelihood of accidents on the other hand [41]. Furthermore, subway drivers’ reactions have been directly correlated to accidents [30], thus paying attention to the high quality of communication with the control room for the subway driver that facilitates the complexity of tasks, and the right place for recovery and rest is of great significance. Major problems that urban subway drivers have reported include vibration, back pain, accident stress, noise pollution and air conditioning that may affect seat and vibration conditions, and the results of this study are also related to depression. Noise pollution and air conditioning have also been addressed among the problems of subway drivers, which has had an impact on depression among underground subway drivers [42].

The safety performance of a transport system depends on the performance of its employees, and their dissatisfaction is associated with the safety performance of the transport system. Much of the dissatisfaction among employees can be expressed and removed, but there are also cases that are hidden factors affecting job satisfaction among employees. According to the result of the satisfaction model, the quality of their resting place is the most important factor affecting the metro drivers’ satisfaction, which is not directly expressed by drivers, and less attention has been paid to it.

For the satisfaction model, only hardware satisfaction from the resting place, quality of communication with the control room, driver’s seat position and noise pollution were significant. But for the depression model, as predicted, psychological and demographic factors such as stress rate from an accident occurrence, fatigue and work experience were also affected by depression, and satisfaction parameters related to loneliness in the cabin and at a relatively high rate. The effect is that the resting place variable is in common with the satisfaction model. Therefore, the likelihood of depression other than psychological conditions can affect certain items of satisfaction. It will therefore be much easier and faster to provide subway driver satisfaction with regard to its impact on hardware variables. However, in relation to depression, which is an important and complex social problem, it is influenced by more and more varied parameters that require more work to improve the situation, which may also lead to overall satisfaction.

It is clear that much of the dissatisfaction and problems can be removed by taking some measures. In addition, it is more cost-effective to reduce the risk of depression among metro drivers. Also, the presence of psychology and psychotherapy units in this transport system will be very helpful. For futures studies, it would be interesting to examine the depression among the participants of this study after removing some of the sources of dissatisfaction and problems and to evaluate the effects of improved conditions on the mental health of metro drivers.

Acknowledgements The authors wish to extend their thanks to the Tehran Metro Transit Company in order to collect and complete the questionnaires.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Author Contributions Conceptualization: KA, SAS, MHA; methodology: SAS, KA; formal analysis and investigation: SAS, KA, MHA; writing/original draft preparation: SAS; writing/reviewing and editing: SAS, KA; data curation: MHA; supervision: KA.

Funding Not applicable.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Jansson A, Olsson E, Kecklund L (2005) Acting or reacting? A cognitive work analysis approach to the train driver task. Rail human factors support integr railway, 1st edn. Routledge, London, pp 41–49
2. Karvonen H, Aaltonen I, Wahlström M, Salo L, Savioja P, Norros L (2011) Hidden roles of the train driver: a challenge for metro automation. Interact Comput 23(4):289–298
3. Cabon P, Coblenz A, Molland R, Fouillot J (1993) Human vigilance in railway and long-haul flight operation. Ergonomics 36(9):1019–1033
4. Sussman D, Coplen M (2000) Fatigue and alertness in the United States railroad industry part I: the nature of the problem. Transp Res F Traffic Psychol Behav 3(4):211–220
5. Warm JS, Parasuraman R, Matthews G (2008) Vigilance requires hard mental work and is stressful. Hum Factors 50(3):433–441
6. Langner R, Eickhoff SB (2013) Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol Bull 139(4):870
7. Edkins GD, Pollock CM (1997) The influence of sustained attention on railway accidents. Accid Anal Prev 29(4):533–539
8. Kecklund G, Akerstedt T, Ingge M, Soderstrom M (1999) Train drivers’ working conditions and their impact on safety, stress and sleepiness: a literature review, analyses of accidents and schedules. National Institute for Psychosocial Factors and Health Stress Research Report, p 288
9. Young MS, Stanton NA, Walker GH (2006) In loco intellegentia: human factors for the future European train driver. Int J Ind Syst Eng 1(4):485–501
