Student performance analysis for better career

[1]Monisha.T.P,[2]Dr.B.Sreepathi

[1]M.Tech 4th Sem, Dept. of CSE, RYMEC, Bellary,India;[2] Head of Dept., Dept of ISE, RYMEC, Bellary ,India.
Email: monishagvictor@gmail.com

ABSTRACT

A college student's life can be primarily categorized into domains such as education, health, social and other activities which may include daily chores and travelling time. Time management is crucial for every student. A self-realisation of one's daily time expenditure in various domains is therefore essential to maximize one's effective output. This paper presents how a mobile application using Fuzzy Logic and Global Positioning System (GPS) analyses a student's lifestyle and provides recommendations and suggestions based on the results.

Keywords: Fuzzy Logic, GPS, Android Application

1 INTRODUCTION

A college student's life is multidimensional. Students are expected to be academically excellent, physically fit and socially active along with managing their daily chores and pursuing their fields of interest. This structure would not only help students to engage all activities but also help them live a balanced life. This practice would eventually help them make better career choices on the basis of their interests. For such a practice one needs to invest a threshold amount of time and effort in all the activities. However only a certain number of students are involved and excel in such a practice. In recent times various student related issues have been addressed by researchers using fuzzy logic. Patel et al. [1] have evaluated student's academic performance considering various factors such as attendance, internal exam, lab assignments, and team- work evaluation. Chrysafiadi and Virvou [2] have developed a fuzzy logic system which understands the forgetting process of a student. Ingoley and Bakal [3] have discarded the traditional methodology of assessment of student performance by also considering personal factors such as stress and accepting the fact that the evaluating system can be non-transparent. Gokmen et al. [4] have made a fuzzy evaluation system which helps to evaluate students on the basis of their performance and the type of examinations by setting up an assessment criterion before an examination. Hameed and Sorensen [5] have developed a reliable and robust system using Gaussian membership functions for student evaluation. Xu et al. [6] have personalized the web-based educational system with respect to learning materials, quiz and advices achieving effectiveness in learning. Huapaya [7] has developed fuzzy student diagnosis model to help teachers evaluate students by providing a high degree of flexibility.

A. About Fuzzy Logic

Over the past three decades, fuzzy logic is widely used in all problem-solving domains. One of the reasons for such instantaneous growth since its inception is its usability across all sectors be it Dynamic Programming, Process Control or Optimization. Fuzzy logic discards the theory of 'Absolute Truth' and instead proposes a new theory of 'Partial Truth', also referred as degree of membership

B. Problem Formulation

The problem can be divided into three major parts:

Data Collection: Using GPS and Google Places API, data collection of all the locations visited and time spent at each location by the user.

Fuzzification: Fuzzification is the crisp input and calculate the values of corresponding membership functions.

Defuzzification: Set up a fuzzy inference system based on certain rules and then return recommendations and suggestions

.

2 WORKING PRINCIPLE

A. Data Collection

A college student is carrying his/her smart phone almost everywhere. Hence using the GPS extraction of his/her position throughout the day is possible. In the application and testing of this paper, the mobile application was developed on Android while the point of interest was extracted using Google Maps API by querying the user's location extracted from GPS.

Let **X** be the set of all tags defined as **X** = *x* is a tag. Analysing the way, a person lives is governed by many parameters, but in a typical student's life we are mainly concerned about one's health, education, leisure and social life. However, a person also invests certain amount of time which fails to fall under these categories. An example of this would be travelling time. Activities like these falls under the *other* category. Now let *S*, *L*, *H*, *W*, *O* be subsets of **X** defined as

$S = \{x \mid x \in \mathbf{X}, x =$ social and x home, work$\}$

$L = \{x \mid x \in \mathbf{X}, x =$ leisure and x f= home, work$\}$

$H = \{x \mid x \in \mathbf{X}, x =$ health and x f= home, work$\}$

$W = \{x \mid x \in \mathbf{X}, x =$ work and x home$\}$

$O = \{x \mid x \in \mathbf{X}$ and $x \not\in S \cup L \cup H \cup W\}$

A *tag x* might belong to one or more of the sets *S*, *L*, *H*, *W*. For example, a person might visit an Amusement Park. In this case the person's *social* and *leisure* purposes are fulfilled. Using this categorization technique, we can extract one's location and time spent at each *tag* for the entire day. TABLE I lists down some locations and their possible purpose of visits. The locations mentioned are basically *tags* other than homeand work.

TABLE I: Sample Locations and Purposes

Location	**Purpose**
Cafe, Restaurant	Going out with friends and family.
Supermarket, Gas Station	Chores
Gym, Ground, Hospital	Exercise or Health Treatment
Cinema Hall, Spa	Leisure and Relaxation
Bank, Business Associates	Work

Weighing criteria: For a given purpose, different locations would have different amount of productivity and impact.
For example, hospital and gym both fall under the *health* category. However, one visits a gym to increase his physical activity and hence visiting a gym has a positive impact on one's health. However, one visits a hospital if he/she has fallen sick. Hence, visiting a hospital has a negative impact on one's health. So, we have to handle these two situations differently.

A function *Y* is defined as *Y:* **X** R such that *Y* (*x*) for every *x* denotes the time spent at the location *tag x*. For example, let *x* = gym. Say *Y* (*x*) = 0.5. This implies a person has spent 30 minutes at a *gym* in the entire day. The unit of time is set in hours throughout this paper.

A function Z_S *is* defined as Z_S*: S* [100, 100] such that ZS(*x*) for every *x S* denotes the intensity of the *tag x* with respect to the *social* category. Similarly, one can define Z_H, Z_L, Z_W, and Z_O for the *health*, *leisure*, *work*, and *other* categories respectively. The range [100,100] is chosen for normalization purposes. For example, let *x* = gym. Say Z_H (*x*) = 50 > 0 as gym has a positive health impact. Let *y* = hospital, then Z_H (*y*) = 20 < 0 as hospital has a negative health impact. However, $Z_L(x) = Z_L(y) = 0$ as both *x* and *y* don't contribute to the *leisure* category. Also note that if a *tag t* belongs to two different categories, then its weightage in both the categories cannot be 0.

For both *Y* and *Z* don't include the *home* tag as it is a special case. This is explained later.

Assigning weights: One is free to assign the weights independently. However, for better results, one can assign weights by conducting a survey to understand how appropriately a location *tag* fulfils the purpose of a category. For instance, consider the *health* category. In the survey a sample population was asked to rank every *x H* in an order of fulfilment of their positive *health* benefits. Consider the following survey with

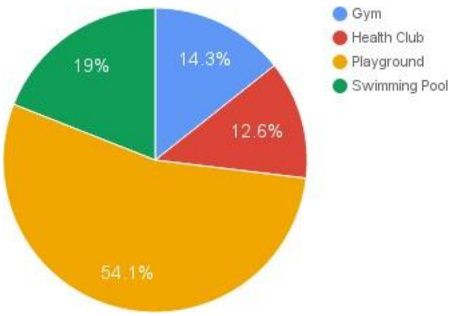


Fig. 1: Survey for *positive* health weights

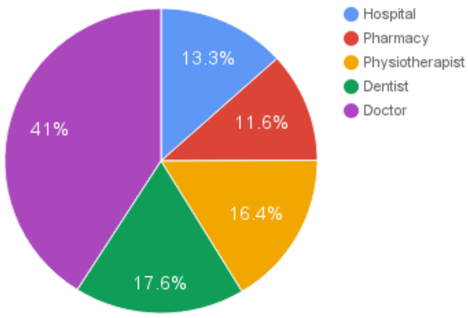


Fig. 2: Survey for negative health weights

H = {gym, playground, swimming_pool, health_club, hospital, pharmacy, physiotherapist, dentist, doctor}

Fig. 1 shows a survey for determining *positive* weights in the *health* category. As 54.1% people taking the survey voted *playground* as their maximum positive benefit from the *health* category, the corresponding weight for x = playground is computed as $Z_H(x) = \frac{54.1}{} \times 100 = 54.1$.

Fig. 2 shows a survey for determining *negative* weights in the *health* category. As 41% people taking the survey voted *doctor* as their maximum non fulfilment from the *health* category, the corresponding weight for x = doctor is computed as $Z_H(x) = -\frac{41}{} \times 100 = -41$.

Home tag: The time spent at the *home* location might not be entirely used for rest and leisure purpose only. One might practice yoga at one's home and the equivalent time should be added to the *health* category. Let τ denote the total time spent at *home*. And τ_H, τ_W, τ_L, τ_O, τ_S denote the equivalent time in respective categories. This time is taken as user input through the mobile application. For better results a random push notification system is used to learn the characteristics of the user. The home tag will be associated with weights ξ_H,ξ_S, ξ_L, ξ_W,ξ_O which denote the intensity of the tags at *home*. For instance, $\xi_W = 30$ and Z_W (office) = 50 > 30 as working at *home* might not be as productive as working at *office*.

B. Fuzzification

Fuzzification of time: Consider a person *p*. Suppose *p visits* tags $\{x_1, x_2, \ldots\ldots, x_n\}$

with the time spent at these locations denoted by $\{Y(x_1), Y(x_2), \ldots, Y(x_n)\}$. Let K_H, K_L,K_S,K_W,K_O denote the overall time spent in *health, leisure, social, work, and other* categories respectively. Then

$$K_H = \sum_{x_1 \in H} Y(x_i) + \tau H$$

Similarly, K_L,K_S,K_W,K_O are defined.

We define the following fuzzy sets for all the categories. These sets define the type of lifestyle of a person is living in each category. Here *leisure* also includes rest.

health = {unfit, fit, proactive}

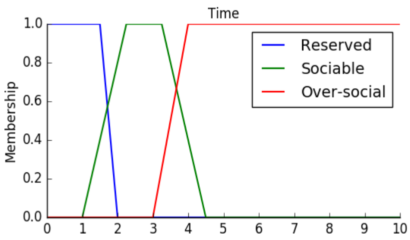
leisure = {hectic, ideal, lazy}

social = {reserved, sociable, over_social}

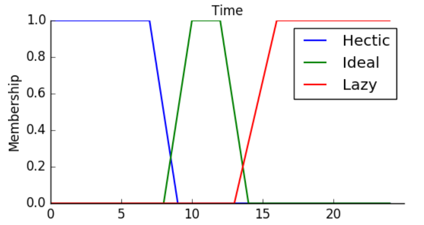
work = {lethargic, hardworking, industrious}

others = {non-productive, productive}

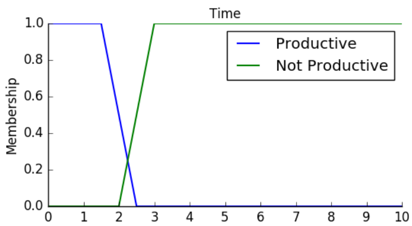
The membership functions for these fuzzy sets are con- structed by conducting a survey on a sample population. The data from the survey can be approximated by using quantile range and trapezoidal membership functions. How- ever, one can use various other techniques to plot membership functions. For instance, in a sample survey the hours spent by fit students in the health *category* were: 0.45, 1.25, 2, 2.25, 2.5, 2.5, 2.75, 2.75, 3, 4, 4.25. So, with respect to the inter quantile range Q_1 = 2, Q_2 = 2.5, Q_3 = 3, inf = 0.45, and sup = 4.25. The trapezoidal membership function for the linguistic term "fit" under the *health* category using these values.



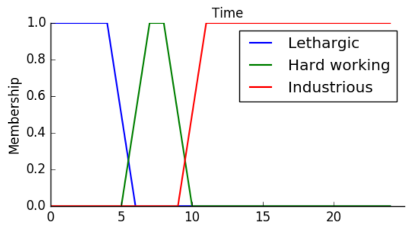
a) Membership function of *social* category



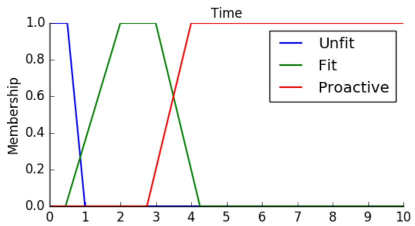
b) Membership function of leisure category



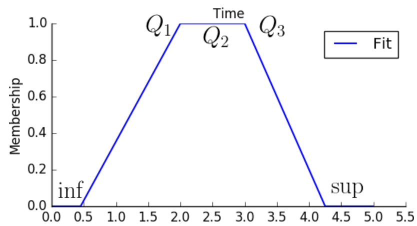
c) Membership function of *another* category



d) Membership function of *work* category



e) Membership function of health category



f) Membership function of the linguistic term *fit*

Fig. 4 shows the membership functions for each linguistic of all categories.

Fuzzification of score: Not only the time spent at a location is important but also how the time is spent is important too. This effective utilization of time is denoted by a *score* MS, M_L, M_O, M_W, and M_H for the respective categories. The score for the *social* category is calculated as follows

$$M_S = \sum_{x \in S} Y(x) Z_S(x) + \tau_s \mathcal{C}s$$

Similarly, other scores can be defined. The fuzzy set of linguistic terms "low_score", "ideal_score" and "high_score" define the fuzzy scores in each category. The membership function of these sets in all categories is calculated similar to the fuzzy time membership functions by conducting a survey. For instance, a survey conducted on a sample of fit students is shown in TABLE II. Hence inf = 11.25, Q_1 = 29.75, Q_3 = 42, sup = 50. Accordingly, the membership function for the linguistic term "ideal_score" under the health category is shown in Fig. 5. Similarly, one can plot the membership functions for the entire fuzzy set across all categories.

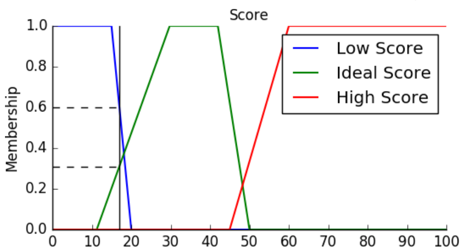


Fig.5: Membership function for "ideal score" under *health* category

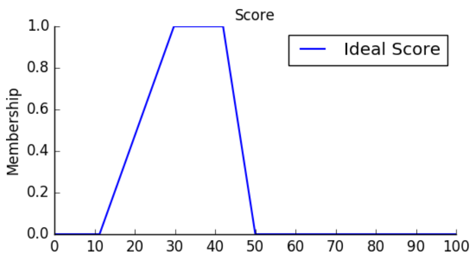


Fig.6: Calculation of membership values

Given the input data, $Y(x)$, $Z_i(x)$, τ_i, and ξ_i where $xi = S, L, O, W, H$, one calculates corresponding K_i and M_i. Using surveys, the membership functions for all linguistic terms in all categories for both fuzzification of time and score can be determined. Hence the membership value of K_i and M_i in the respective categories for all the linguistic terms can be determined. Let R_1, R_2, R_N be a set of recommendations. Now every R_k $(1 \leq k \leq N)$ will be dependent on a set of linguistic terms. For example, a recommendation R = *"All work and no play make Jack a dull boy."* will be outputted if a person is spending too much time and effort in work and less in his leisure and social life. That is, he/she has a "industrious" work life with a high work score and has a "reserved" social life with a low social score and a "hectic" life with respect to leisure with a low score. So, attributes of R can be represented as {K_W = "industrious", M_W = "high_score", K_S = "reserved", M_S = "low score", K_L "hectic", M_L = "low_score"}

Let R_k be a recommendation with attributes $\{a_1, a_2, a_3, a_n\}$. Here each a_j $(1 \leq j \leq n)$ is a combination of score/time with respect to a linguistic term of a category. Hence as shown previously one can calculate its membership value for each attribute. Here n can vary for each R_1. For instance, Fig. 6 shows the membership functions of M_H . Let a_1, a_2, a_3 be the following attributes

$a_1 = M_H$: *low_score*
$a_2 = M_H$: *ideal_score*
$a_3 = M_H$: *high_score*

Hence μ_1, μ_2, μ_3 for $M_H = 17$ as seen from Fig. 6 will be 0.6, 0.310, 0.0 respectively. Using equal weighing criteria for each a_j, we can calculate a score of each recommendation $\rho(R_k)$ defined as

$$\rho(R_k) = \frac{1}{n}\sum_{j=1}^{n}\mu_j$$

Now, using the *most probable criterion* the recommendation with the maximum score value $\rho(R_k)$ will be displayed as output.

3 EXPERIMENT

A survey conducted in IIT Kharagpur was conducted to determine all the membership functions for all linguistic terms across all the categories. Some of the membership functions are shown in this paper. The mobile application was installed on the student's phone and the results were analyzed. A random student was picked and his data for the day was analyzed. TABLE III shows the *tags* he visited throughout the day and their corresponding time and weights. The score for each *tag* is also enumerated. TABLE IV shows the total time and score across all the categories. The recommendations in the set R where $R = \{R_1, R_2, R_3, R_4\}$ were considered.

R_1 = *"Catch up a movie this evening."*

R_2 = *"Work is worship."*

R_3 = *"Family matters."*

R_4 = *"Hit the gym."*

The attributes of R is shown in TABLE V. The membership values for each attribute is shown in TABLE VI and the corresponding score of each recommendation is also enumerated. As $\rho(R_1)$ is maximum the mobile application recommended the student to *"Catch up a movie this evening."*

TABLE III: Experiment Data

Tag	Time	Weight	Score
university	$Y(x) = 6$	$ZW(x) = 50$	300
library	$Y(x) = 4$	$ZW(x) = 20$	80
home	$\tau W = 2$	$\xi W = 30$	60
	$\tau S = 0.5$	$\xi S = 30$	15
	$\tau H = 0.5$	$\xi H = 20$	10
	$\tau L = 6.5$	$\xi S = 30$	195

	$\tau_O = 1$	$\xi_O = 10$	10
cafe	$Y(x) = 1$	$Z_S(x) = 20$	20
supermarket	$Y(x) = 1$	$Z_O(x) = 9$	9
grocery	$Y(x) = 0.5$	$Z_O(x) = 10$	5
travel	$Y(x) = 1$	$Z_O(x) = 15$	15

TABLE IV: Experiment Calculations of total time and score for each category

Total time	**Total score**
$K_S = 1.5$	$M_S = 35$
$K_L = 6.5$	$M_L = 195$
$K_O = 3.5$	$M_O = 39$
$K_W = 12$	$M_W = 440$
$K_H = 0.5$	$M_H = 10$

TABLE V: Experiment Recommendation attributes

Recommendation	**Attributes**
R_1	{K_L = "hectic", M_L = "less_score", K_W = "industrious", M_W = "high_score"}
R_2	{K_W = "lethargic", M_W = "less_score", K_L = "lazy", M_L = "high_score"}
R_3	{K_S = "reserved", M_S = "less_score"}
R_4	{K_H = "unfit", M_H = "less_score"}

TABLE VI: Experiment Recommendation score calculation

Recommendation	**Membership Values** (μ_j)	**Score** ($\rho(R_k)$)
R_1	{1.0, 0.8, 1.0, 1.0}	0.95
R_2	{0.0, 0.0, 0.0, 0.0}	0.0
R_3	{1.0, 0.7}	0.85
R_4	{1.0, 0.8}	0.9

4 CONCLUSION

From our proposed work we can conclude that the analysis of the activities of the students is very much important for a bright career. From our analysis we have observed that student's day to day activities and their involvement in various in campus and out campus activities vary and their performance during the exam are interlinked. The proposed work can predict and give recommendations to the students based on fuzzification. The weight matric of each and every activity are calculated based on that ten intelligent scores has been assigned to every student. Students performance are vandalized during the exam time and they depend on the time which they spend on various places and the activities which they perform. Based on this the students are classified into three categories namely cream student, normal student and average student.

ACKNOWLEDGMENT

The authors would like to thank Professor Sudhirkumar Barai of the Civil Engineering Department, IIT Kharagpur for his continued and unconditional guidance. An exemplary teacher and a magnificent person, we consider ourselves lucky to have been taught by him and to have worked under his supervision. Without his course, "Soft Computing Tools in Engineering", and his co-operation the preparation of this paper would not have been possible.

REFERENCES

[1] S. Patel, P. Sajja, and A. Patel, "Fuzzy logic based expert system for student's performance evaluation in data grid environment," *International Journal of Scientific & Engineering Research*, vol. 5, no. 1, 2014.

[2] K. Chrysafiadi and M. Virvou, *Fuzzy Logic in Student Modeling*, pp. 25–60. Cham: Springer International Publishing, 2015.

[3] S. Ingoley and J. Bakal, “Evaluation of student performance in laboratory applications using fuzzy logic,” *International Journal on Advanced Com- puter Engineering and Communication Technologies*, vol. 1, pp. 2278 – 5140, 2012.

[4] G. Gokmen, T. C. Akinci, M. Tektas, N. Onat, G. Kocyigit, and N. Tektas, “Evaluation of student performance in laboratory applications using fuzzy logic,” *Procedia - Social and Behavioral Sciences*, vol. 2, no. 2, pp. 902 – 909, 2010.

[5] I. A. Hameed and C. G. Sorensen, *Fuzzy Systems in Education: A More Reliable System for Student Evaluation*, ch. 1. InTech, 2010.

[6] D. Xu, H. Wang, and K. Su, “Intelligent student profiling with fuzzy models,” in *Proceedings of the 35th Annual Hawaii International Con- ference on System Sciences (HICSS’02)-Volume 3 - Volume 3*, HICSS ’02, (Washington, DC, USA), pp. 81.2–, IEEE Computer Society, 2002.

[7] C. R. Huapaya, “Proposal of fuzzy logic-based studentslearning assess- ment model,” *XVIII Congreso Argentino de Ciencias de la Computacin*, 2012.

[8] L. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338 – 353, 1965.

[9] G. J. Klir and B. Yuan, eds., *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh*. River Edge, NJ, USA: World Scientific Publishing Co., Inc., 1996.