

XPEN – A Voice Powered Expense Tracker Full Stack Web Application

Ms. Kritika¹, Dr. Himani², Ms. Shikha³

^{1,2,3}Assistant Professor

Department of Computer Science and Engineering (CSE)

Bhagwan Parshuram Institute of Technology

Delhi, India

Abstract— Budgeting is an integral part of society. From the beginning of civilization humans have been selling and purchasing goods. Since then, it has become an essential and irreplaceable part of our lives. Many of us people live in middle class families. We have to track our daily budget to avoid the debt. Tracking our expenses daily can save our amount, but it can also help us set financial goals for the future. If we know exactly where our amount is going every month, we can easily see where some cutbacks and compromises can be made. Most of us have a fixed income and we get it on time (i.e., daily, monthly, annual, etc.). In addition, everyone follows a strict budget of spending. Generally, the budget is assembled according to category. Categories vary, for example, food, entertainment, transportation, education, health, clothing, and so on. However, spending is limited to budget revenue. For this reason, we need to keep track of our expenses so that they do not exceed our budget. People often track expenses using pen and paper methods or take notes on a mobile phone or a computer. These processes of storing expense require further computations and processing for these data to be used as a trackable record. In this work, we are proposing an automated system named XPEN to store and calculate these data. XPEN is a web application which aims to assist a user in managing personal or family finances by offering not only a basic expense check but also a brief analysis of incomes and expenses. All operations are performed in private through a personal account. Voice capabilities are also added to this application with the help of which users can save their expense by just speaking. The application solves all of the problems that occur with old ways of entering data. It uses the voice to traverse the app. This

application filters the keywords from the user’s voice and saves the amount and description for further processing. All transactions of the user are accumulated to the daily, weekly and monthly total sum and visualized as a histogram. The user then gets a better understanding of the weekly, monthly and yearly fluctuation of incomes and expenses. XPEN can manage daily expenses much faster than any other traditional app in the market which takes manual input. Overall, this is a smart automated solution for tracking expenses.

Keywords—Expenditure, Track Expense, VUI, Voice Recognition, Voice Interface, User Interface, Automatic Speech Recognition

I. INTRODUCTION

From the beginning of human civilization, people have exchanged their fortune with each other for buying or selling goods. It has become a crucial and unchangeable part of our daily life since then. Most of us have a fixed income and we get it in a timely basis (i.e. daily, monthly, yearly etc.). Moreover, everyone follows a strict budget of expense. Generally, the budget is assembled as per category. The categories are distinct, for example, food, entertainment, transportation, education, healthcare, clothing etc. However, the budget of expense is restricted to the income. For that reason, we need to track our expense so that it doesn’t exceed our budget. In old days, people used to track their expense manually i.e., using pen and paper system which takes a lot of effort and time.

Nowadays, the availability of electronic devices like smartphones, computers have made our life a lot easier and faster. We can use computers to track our daily expense by using the online and offline software available. But the computer is not accessible all the time. The smart solution to the problem is to use smartphones. Nearly 44% of the world population use smartphones [1]. Smartphones have become an irreplaceable part of our daily lives as they are always accessible on the go. There are some existing applications that can track daily expense [2]–[5]. These applications use a manual input system from the keyboard which is tiresome and time-consuming. To meet the challenge of avoiding manual input, we are proposing a smart method of doing the same work but in a more automated and efficient way which takes less time.

In the proposed approach, users can save their expense simply by just speaking. To capture the user’s voice, the ‘sounddevice’ library for python has been used. The ‘wavio’ package in python has been used for file I/O to save the audio in wav format. The ‘wav’ format provides lossless

compression for the audio which means that no detail from the input voice is left out. The python’s ‘Speech Recognizer’ library is used for the conversion of received audio file into text. This library takes in a ‘wav’ file as an input and predicts the spoken text/ transcript as an output. The Processed voice command after conversion to text needs to be interpreted as an executable command/sequence. This mapping will be done by wit.ai. The ‘wit.ai’ model is based on a probabilistic language model developed on artificial neural network.

The sections are structured as follows: in Section 2, the background analysis is covered. It includes some of the popular applications which are available. Section 3 defines our approach, Section 4 is the implementation and Section 5 includes a conclusion.

II. RELATED WORKS

Expense management apps are very common in the application market. Many of them offer exciting features. Different apps have taken different approaches to manage the daily expense.

Daily Expense 3 [2] is a system that can track income and expense and classify them into categories. The application shows reports grouped by periods. Users can also schedule their recurring records. The application also creates a backup of their records to restore information if necessary.

AndroMoney [5] supports multiple accounts to manage expense and income. It uses cloud storage so that the data is safe. Users can set a budget for the expense and the app will notify if they exceed the budget. It provides a number pad to calculate any record. It generates trend, pie and bar charts for cash flow.

Monefy [3] - money manager is an expense management application that has an intuitive interface. It can store a record faster than the others mentioned above. It provides widgets to enhance ease of access. It provides default categories and option to add customized categories. It also provides an onscreen calculator to calculate the expense.

Expense Manager [12] is a feature-rich application, which is interactive and well balanced. Besides tracking user’s income and expense it also saves the picture of a receipt. It also tracks tax, mileage, and debts. It provides some convenient tools such as currency converter, regular calculator, tip calculator, loan calculator, credit card payoff calculator etc.

The next application which is being used to track expense is Money Lover [4]. This application manages expense category wise. It also manages income, debts, and loans. Users can set events and save plans on this app. It saves the receipt. It also provides a currency converter and a calculator.

In our work, we have developed a system which has the functionalities of a traditional expense management web applications along with the voice capability to take automatic input. The system is more interactive and operational than all the other solutions mentioned above.

III. PROPOSED SYSTEM

The project is a MERN stack application, The front-end will be developed in HTML5, CSS3, and React.js while the back- end will be created in Express and Node.js. The data will be stored in MongoDB Atlas and users can only access their own data not others.

A. User Story

After login, the user can access the following key features:

- (i) *Add/ Edit/ Delete a transaction:* The user can create a new income or expense transaction which can be deleted or modified after the creation.
- (ii) *Check recent transactions:* The user will see all transactions took place in the last three days and the sum of income, expense and balance in the personal dashboard.
- (iii) *Check history transactions:* The user can also check the list of previous transactions along with the sum and ratio of income and expense in a day, a week, a month and a year.

(iv) *Get an overview of incomes and expenses:* All transactions of the user are accumulated to the daily, weekly and monthly total sum and visualized as a histogram. The user then gets a better understanding of the weekly, monthly and yearly fluctuation of incomes and expenses.

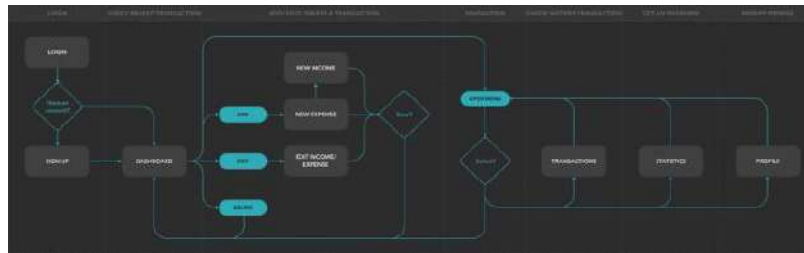


Fig. 1. User Stories

B. Web Development

The project is a MERN stack application, deployed on Heroku. Since personal financial records are sensitive information, it is necessary to keep personal data in private. Therefore, user authentication was required and verified with JSON Web Token (JWT). By doing so, a user can login and log out with a personal account.

(i) Front-end

The front-end was built with responsive modern technologies (HTML5, CSS3, React.js). The data analysis was achieved by visualizing transaction data with D3.js.

(ii) Back-end

A REST API for the project was created with Express, Node.js while the data was stored in MongoDB Atlas, a NoSQL and document-oriented database in the cloud. Then, a user can perform CRUD operations through this API such as create, read, update, and delete data.

C. VUI System

The architecture of the complete VUI system consists of modules: The details of various modules are described below:

A. Speech Recording

To capture the user's voice, the 'sounddevice' library for python has been used. The library provides recording of audio signals via single channel and dual channel mode. It also allows the users to select the primary recording device. This library allows the storage of sound in numpy arrays as digital sampled signal at the given frequency. The Numpy arrays are thus processed to get the text from the audio signal.

B. Audio File Handling

The 'wavio' package in python has been used for file I/O to save the audio in wav format. The 'wav' format provides lossless compression for the audio which means that no detail from the input voice is left out. The wavio package stores the output sound signal as an uncompressed 8 bit pcm format. The number of bits in the PCM format denote the depth at which the sound has been captured. Higher bits implies higher detail. However the 8 bit format is the best one to choose due to its relatively less space complexity and good enough quality for speech recognition.

C. Speech To Text API

The python's 'Speech Recognizer' library is used for the conversion of received audio file into text. This library takes in a 'wav' file as an input and predicts the spoken text/ transcript as an output. The SpeechRecognizer library can be used in 2 variants. online mode and offline mode. As ours is a web application, we are using online mode. This mode uses various online speech to text tools to convert speech to text. One API that is used is the Google's web search speech api. The advantage of this mode is that the major players in the speech to text grounds can be utilized and the speech recognition is accurate.

D. Natural Language Processing API

The Processed voice command after conversion to text needs to be interpreted as an executable command/sequence. This mapping will be done by wit.ai. Every voice command that is given to the computer has 2 parts:

- 1) Intent: This is the objective of the user to complete a specific work/task that is given to the computer.
- 2) Parameter: The parameter is the object on which the intent has to be performed.

The two part bifurcation of the NLP model make its highly easy to understand and interpretable. The 'wit.ai' model is based on a probabilistic language model developed on artificial neural network. This model is based on the theory of n-grams for probability prediction i.e. it can predict the probability of a given word in a sentence or a statement by usage of conditional probability on the statement from the statistics of the text corpus. The Model can predict the probabilities of intents and parameters from sentences similar to the trained sentence. Along with the use of probability, the model also makes use of text labelling also known as Named Entity Recognition (hereafter referred to as NER). The NER labels the parts of the text to the pre-defined labels. It is achieved by finding patterns in the text that has been used for the training of the model. For example, consider the sentence Add Expense, the intent identified is Add and the argument identified is Expense.

E. System Programming

The system programming module does the role of executing the actual commands on the host machine. For implementation, following system API operations were added: Add Expense, Add Income, Add Rs. <Amount>, Add in Category Travel, Add in Category Entertainment, Add in Category Food, Add in Category Clothing, Add in Category Health, Add in Category Education.

IV. IMPLEMENTATION

XPEN is divided into seven major sections. Those are Login, Check Recent Transaction, Add/Edit/ Delete a Transaction, Navigation, Check History Transaction, Get an Overview and Modify Profile. This system works as a one tap solution for tracking everyday expense. It also preserves yearly and monthly records. For the availability of the records, users can check their histories to keep track of their expense so that they do not exceed their pre-allocated budget. First of all we broke up the UI in components. We used context api. So, we have global state, global context api, we passed everything to the components. Basically we put all the transactions in the global state, and passed it to the components, then we did all the calculations stuff inside that. The transactions are stored in MongoDB database.

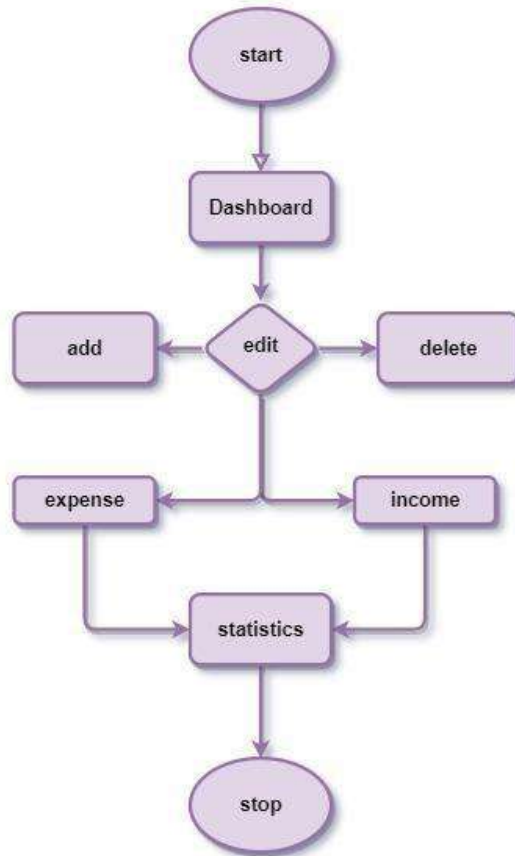


Fig. 2. Flow of UI Components of the App

In this project there is only one user. The user start the program on the user dashboard it show three option add,edit and delete after its calculate and add to database and show final result in statistics and user stop the program.

The screenshot shows a web form titled "Add new transaction". It contains two input fields: "Text" with a placeholder "Enter text..." and "Amount" with a placeholder "Enter amount...". Below the "Amount" field, there is a note: "(negative - expense, positive - income)". At the bottom of the form is a blue button labeled "Add transaction".

Fig. 3. Add Transaction

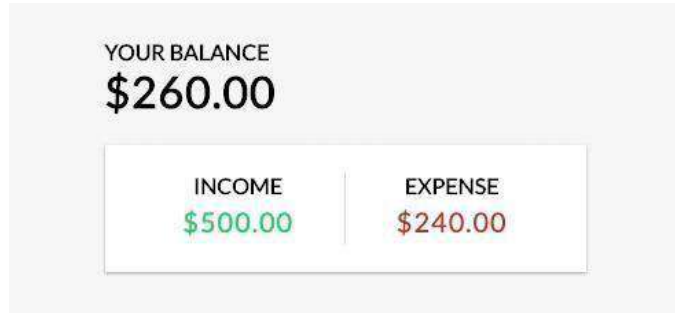
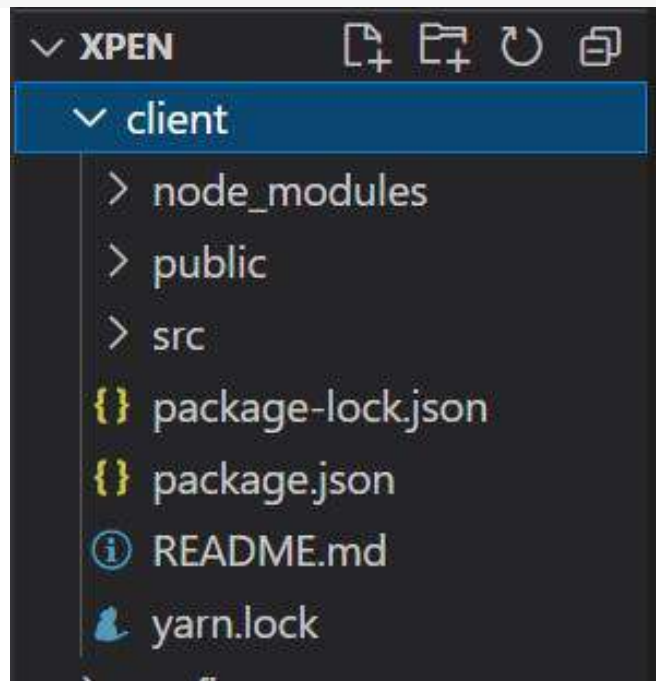


Fig. 4. Balance Section of the App



Fig. 5. History

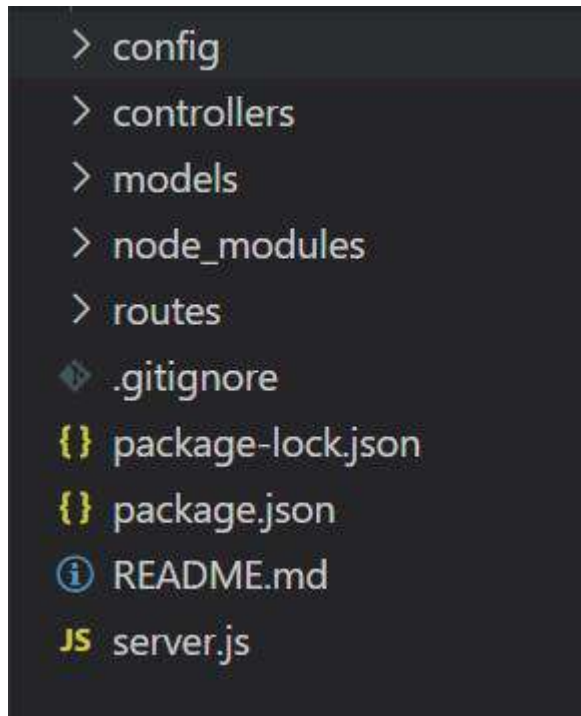
Section of the App



Folder

Structure

Fig. 7. Server the App

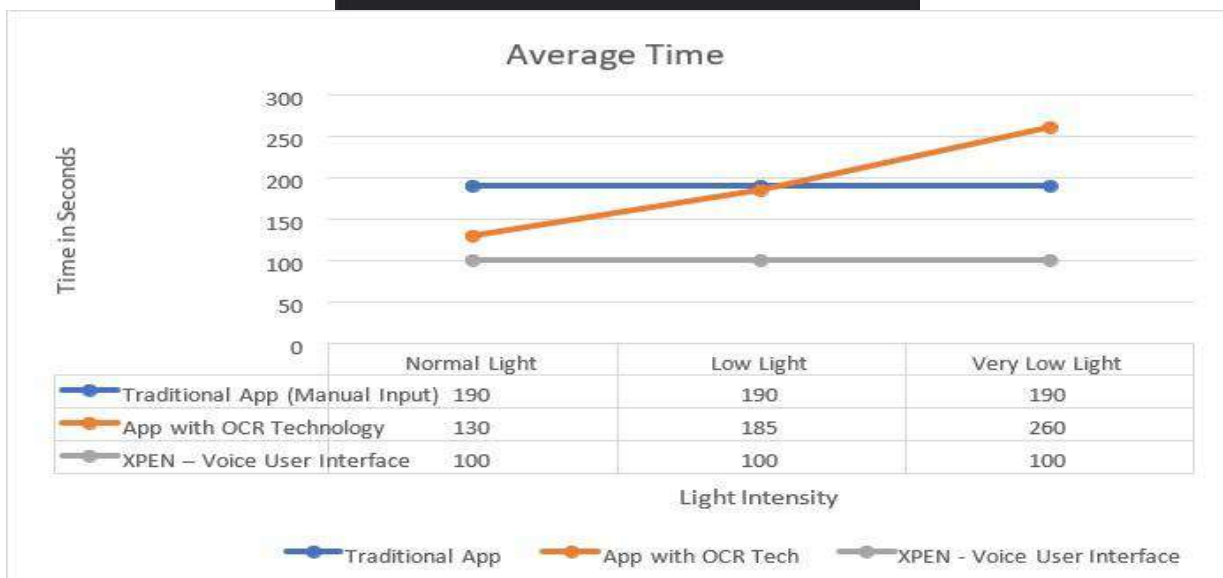


Side of

time taken to add or

V. ANALYSIS

Fig. 8 Average delete a transaction



There are mainly four tasks that can be performed by the user in the XPEN app. There are “add income”, “add expense”,

“delete income” and “delete expense”. The user can perform anyone of these tasks. So, let’s say the user performs anyone of these tasks repeatedly for several number of times. To find the average time taken to complete a task, divide the total time

Fig. 6. Client Side of the App taken to performed the tasks by the total number of tasks performed by the user. There are three series in the graph namely traditional app, app with OCR tech and XPEN. There are three categories namely normal light, low light and very low light.

There is no effect of light intensity on the traditional app and the XPEN [VUI] app. But the performance of app using OCR technology decreases when the intensity of light decreases. The character recognition from the receipt declines in low lights. XPEN performs better than both the traditional app and the app with OCR technology. It takes lesser time to complete all the tasks.

VI. CONCLUSION

In today’s world, time is the most valuable asset because people lack ample of it. People are

obsessed with completing tasks in lesser time and our system is an approach serving this purpose. XPEN can manage daily expense much faster than any other traditional app in the market which takes manual input. Nowadays, the world is leaning towards the one tap solution and our system is one of a kind. After all, automation is the way of future and XPEN can be a step towards it. In the future, the AI technologies of speech recognition and generation is of the most difficulties to break through. With developed AI technologies, VUI can make machines hear and talk humanly, so the convenience and accessibility of UI will come to another leap after GUI. In addition, from the aspect of ethic, new policies ensuring the fairness of access to VUI should be put forward so that people in the society could take advantage of VUI equally.

There are mainly four tasks that can be performed by the user in the XPEN app. There are “add income”, “add expense”, “delete income” and “delete expense”. The user can perform anyone of these tasks. So, let’s say the user performs anyone of these tasks repeatedly for several number of times. To find the average time taken to complete a task, divide the total time taken to performed the tasks by the total number of tasks performed by the user.

Therefore: There is no effect of light intensity on the traditional app and the XPEN [VUI] app.

VII. REFERENCES

- [1] “44% of World Population will Own Smartphones in 2017.” [Online]. Available: <https://www.strategyanalytics.com/strategyanalytics/bl ogs/smart-phones/2016/12/21/44-of-world- populationwill-own-smartphones-in- 2017#.Wq1BcehuY2y>. [Accessed: 17Mar-2018].
- [2] “Daily Expense 3 - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=mic.app. gastosdiarios>. [Accessed: 17-Mar-2018].
- [3] “Monefy - Money Manager - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=com.mo nefy.app.lite>. app performs poorly if the input image from the receipt includes a lot of noises. The app is unable to detect the region of interest automatically. On that note, the user has to set the region of the receipt to have a better result. Again, the performance of the [Accessed: 17-Mar-2018].
- [4] “Money Lover: Budget Planner, Expense Tracker - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=com.bookmark.money> . [Accessed: 17-Mar-2018].
- [5] “AndroMoney (Expense Track) - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=com.kpmoney.android> . [Accessed: 17- Mar-2018].
- [6] “Expense Manager - Apps on Google Play. [Online]. Available: <https://play.google.com/store/apps/details?id=com.expensemanager>. [Accessed: 17-Mar-2018].
- [7] Sabab, S. A., Islam, S. S., Rana, M. J., & Hossain, M. (2018, September). eExpense: A Smart Approach to Track Everyday Expense. 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (ICEEiCT).<https://doi.org/10.1109/ceeict.2018.8628070>
- [8] Yurochkin, D. E., Horoshiy, A. A., & Karpukhin, S.A. (2021, January). Development of an Application for Expense Accounting. 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). <https://doi.org/10.1109/elconrus51938.2021.9396201>
- [9] Dabre, P., Gonsalves, R., Chandvaniya, R., & Nimkar, A. V. (2020, April). A Framework for System Interfacing of Voice User Interface for Personal

Computers. 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA).

<https://doi.org/10.1109/cscita47329.2020.9137809>

[10] Begany, G. M., Sa, N., & Yuan, X. (2015). Factors Affecting User Perception of a Spoken Language vs. Textual Search Interface: A Content Analysis. *Interacting with Computers*, iww029.

<https://doi.org/10.1093/iwc/iww029>

[11] Zhang, X., Lu, C., Yin, J., Xie, H., & Liu, T. (2020).

The Study of Two Novel Speech-Based Selection Techniques in Voice-User Interfaces. *IEEE Access*, 8, 217024–217032.

<https://doi.org/10.1109/access.2020.3041649> [12] 2016 International Conference on Computational

Techniques in Information and Communication Technologies (ICCTICT)

[13] *International Journal of Commerce & Business Studies* Volume-4, Issue-1, January-March, 2016, www.iaster.com ISSN (O) 2347-2847 (P) 2347-8276

[14] 2020 International Conference on Computing and Data

Science (CDS) 978-1-7281-7106-7/20/\$31.00 ©2020 IEEE DOI 10.1109/CDS49703.2020.00045

[15] 2020 International Conference for Emerging

[16] Technology (INCET) Belgaum, India. Jun 5-7, 2020 978-1-7281-6221-8/20/ ©2020 IEEE 1 Authorized licensed use limited to: Cornell University Library. IEEE Xplore. Restrictions apply

T. Kita, C. Nagaoka, N. Hiraoka and M. Dougiamas, "Implementation of Voice User Interfaces to Enhance Users' Activities on Moodle," 2019 4th International Conference on Information Technology (InCIT), 2019, pp. 104-107, doi:10.1109/INCIT.2019.8912086.