# Use-Case

**Money Tracker APP** with Voice Input

- Customers could use EchoBudget at home and outside with the device
- Customers could verbally:
  - Record the spending
  - Modify or remove existing entry
  - Request financial report for a given time range
- **Friendly to visually impaired customers & who wants to record their expenses in a easy way**

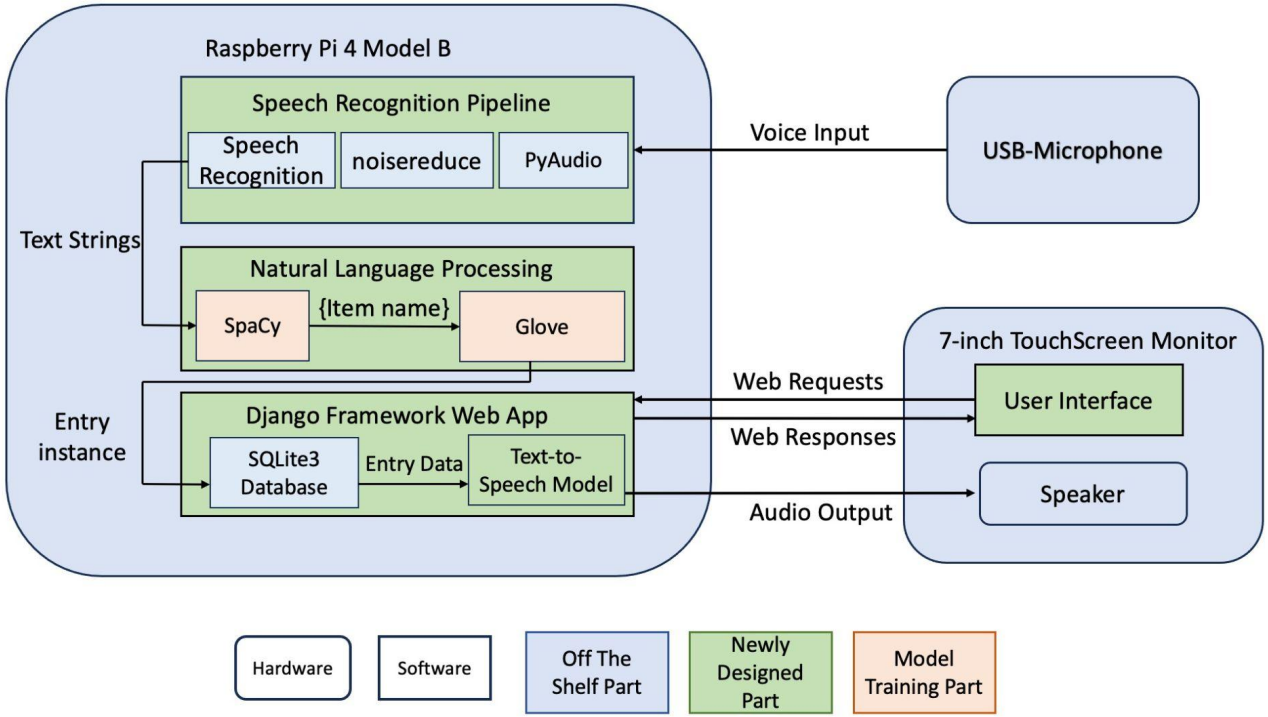
# Use Case Requirements

- **Latency**
  - Handle request and provide response within 4s for each user input
- **Portability**
  - Weight:  <= 500g (weight of iPad Air)
  - 1-hour battery life (with monitor on)
- **Accessibility**
  - Speech recording button size big enough(32% of space) for visually impaired group to click without difficulty
- **Noise Reduction**
  - All functions should work with 90% accuracy under 70 dB environment

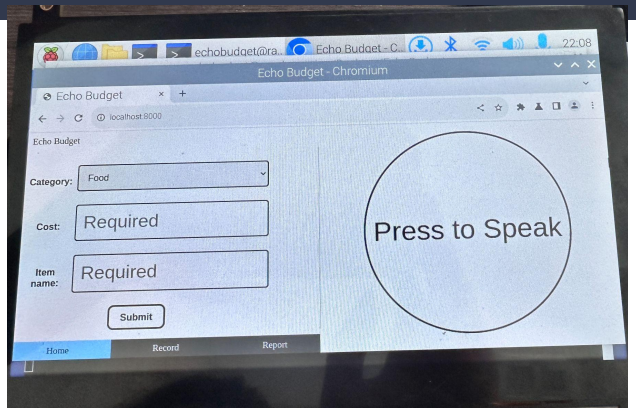
# Design Requirements

- **Speech Recognition**
  - The speech to text accuracy for whole sentence >= 80%
  - The accuracy for important words >= 90%
- **NLP**
  - Accuracy for parse the action/money amount/items name from script >= 95%
- **Item classification**
  - Accuracy for categorization should >= 90% for our given categories
  - Categories: Food, housing, necessities, entertainment, transportation, others
- **Active Response**
  - Corresponding voice response should be provided after all voice commands
  - E.g. “Entering XX for $XX, please confirm” after the user enters a new entry

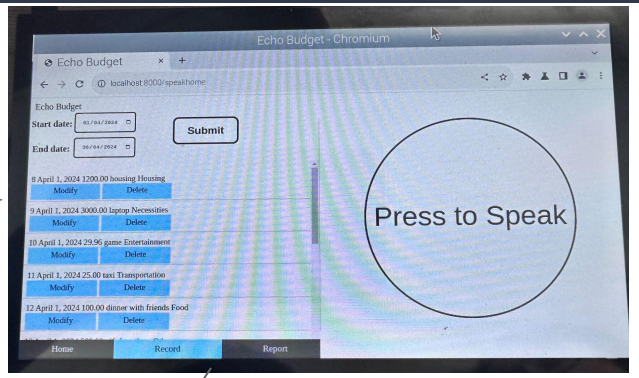
# Solution Approach: Block Diagram



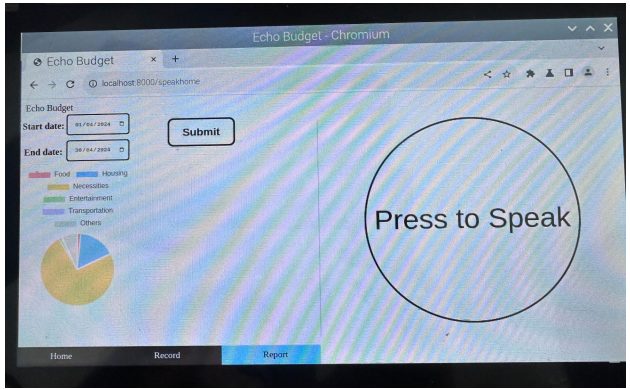
# Complete Solution




"Get Entries"




"Enter cake for 12 dollars"




"Generate Report"

# Testing: Use Case

| Test | Input | Output | Expected Performance |
|---|---|---|---|
| Latency (voice command) | Voice request | Time to render | 4 seconds to render the page |
| Battery life | NA | Power consumption | Consumes less than 50% of power with monitor on for 1 hour |
| Portability | Device | Weight | 500 grams |
| Accessibility | button | Proportion on the screen | 32% of right half of screen |
| Noise reduction | audio | text | 90% accuracy of integral test under 70dB environment |

# Testing: Use Case

| Test | Expected Performance | Actual Performance |
|---|---|---|
| Latency (voice command) | 4 seconds to render the page | Average 4.52 seconds to render the page |
| Battery life | Consumes less than 50% of power with monitor on for 1 hour | Power drops from 100% to 77%, consuming 23% of total power |
| Portability | 500 grams | 500 grams |
| Accessibility | 32% of right half of screen | 35.45% of right half  of screen |
| Noise reduction | 90% accuracy of integral test under 70dB environment | 9 out of 10 commands work as expected, 90% accuracy |

# Testing: Design

| Test | Input | Output | Expected Performance |
|---|---|---|---|
| Audio to text | Digital signal of audio input | Text (list of words) | Less than 20% of word error rate (WER) |
| Text to command (NLP) | Text output from above | Command (verb + params) | Identify verb with 100% accuracy;<br>Identify item name, money, date, and number with 95% accuracy;<br>NLP process takes less than 3s |
| Item classification (Glove) | Item name | Category | 90% of item names correctly classified |
| Voice response | Voice command | Voice output | All audio requests should be assisted with a voice response |

# Testing: Design

| Test | Expected Performance | Actual Performance |
|---|---|---|
| Audio to text | Less than 20% of word error rate (WER) | **98.3% accuracy, edge case exists** |
| Text to command (NLP) | Identify verb with 100% accuracy;<br>Identify item name, money, date, and number with 95% accuracy;<br>NLP process takes less than 3s | Identify **verb, money, date, and number** with **100%** accuracy;<br>Identify **item name** with **96%** accuracy<br>NLP process takes about **2.5s** |
| Item classification (Glove) | 90% of item names correctly classified | 18 out of 20 item names are correctly classified, 90% accuracy |
| Voice response | All audio requests should be assisted with a voice response | The app will read out the content of the response (entries, report, etc.) |

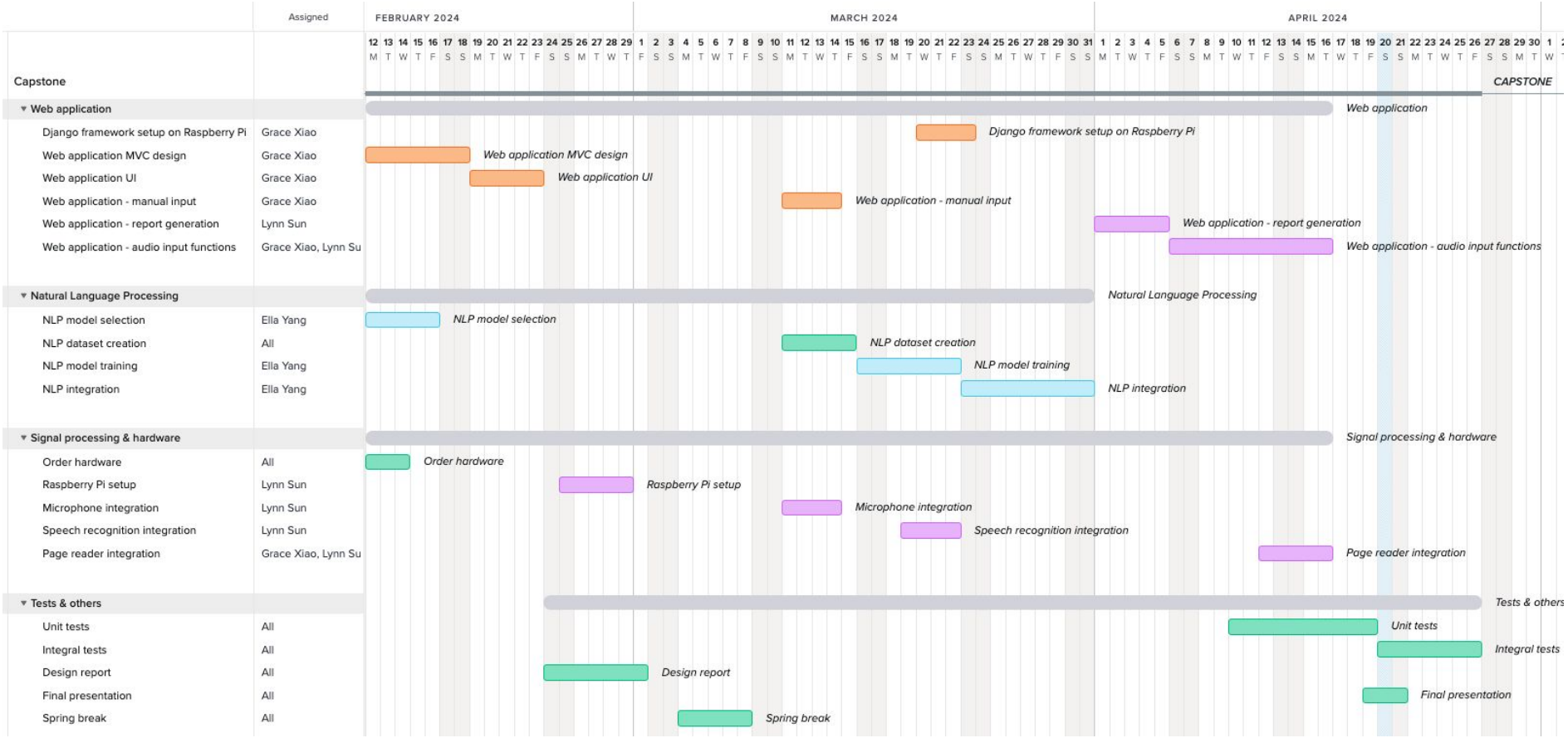
# Integral Testing (planned but not finished)

- User Experience Test
  - 5 volunteers
  - Each volunteer would have **half of the day** to interact with our product **without our interference**
  - Volunteer are expected to give feedback about what they **like and dislike** about the product
- Overall Latency
  - Test different commands (e.g. record/change/delete entry, view entry, generate report, etc.)
  - **Expect <15s** for all commands
- Overall Accuracy
  - **Expect >95%** accuracy for the whole system

# Design Trade-offs

| Train one model for Spacy | **Train two models for Spacy** |
|---|---|
| Require less storage | Increase parsing accuracy<br>Only lengthen the runtime by a little bit |
| Click button to stop | **Automatic speech termination detection** |
| More flexibility on voice input time<br>may cause potential waste if forget to press twice | User does not need to press again<br>Fixed voice inputting time (8 seconds) |
| Allow users to give commands without restriction | **Standardized voice input keywords** |
| More user-friendly, but hard to process | Enhance action performance accuracy |
| Add each item name to Glove model | **Only add item name to model when entry is modified** |
| More accurate, but may overfit | Reduce wait time and storage |

# Schedule


Assigned
FEBRUARY 2024
MARCH 2024
APRIL 2024
Capstone
CAPSTONE
Web application
Django framework setup on Raspberry Pi Grace Xiao
Web application MVC design Grace Xiao
Web application UI Grace Xiao
Web application - manual input Grace Xiao
Web application - report generation Lynn Sun
Web application - audio input functions Grace Xiao, Lynn Su
Natural Language Processing
NLP model selection Ella Yang
NLP dataset creation All
NLP model training Ella Yang
NLP integration Ella Yang
Signal processing & hardware
Order hardware All
Raspberry Pi setup Lynn Sun
Microphone integration Lynn Sun
Speech recognition integration Lynn Sun
Page reader integration Grace Xiao, Lynn Su
Tests & others
Unit tests All
Integral tests All
Design report All
Final presentation All
Spring break All