**Project Work Breakdown Structure**

1. **Initiation**

**1.1 Project objectives**

* Aim to recommend fair prices for items (second hand guitar) based on images.

1. **Project Planning**

**2.1 Timeline**

* Week 4-6 = Data collection & Research methodology
* Week 7-8 = Benchmark baseline model
* Week 8 (March 21) = Midterm Presentation
* Week 9-11 = Model Selection and Improvement
* Week 12-14 = Inference, Analysis & Reporting
* Week 15 (April 30) = FInal Presentation

**2.2 Allocate Tasks**

* Research methodology = Lauri & Vyshnavi
* Benchmarking baseline model = Lauri & Vyshnavi
* Data collection = Xiang & Yan
* Model development and Improvement = Xiang & Yan
* Reporting = All

1. **Literature Review / Research**

**3.1 Review relevant literature on deep learning for price prediction and second-hand instruments**

* Deep Learning for price prediction of second-hand items (Fathalla et al., 20201)
* Vehicle Price Prediction using Visual Features (Yan et al., 20182)

**3.2 Research existing methodologies and approaches in the field**

* Deep Learning for price prediction of second-hand items (Fathalla et al., 20201)
  + The study makes two main contributions:

1. **Price Prediction Model:** The first contribution involves predicting the price of second-hand items using a hybrid CNN-LSTM model that combines images and textual data. The proposed model outperformed the baseline SVM model, with Mean Absolute Error (MAE) scores of 0.07 and 0.09, respectively.
2. **Price Range Forecasting:** The second contribution focuses on improving price prediction accuracy by forecasting the minimum and maximum prices of second-hand items. By combining the item quality score (predicted price) with the forecasted price range, the study enhances the overall price prediction accuracy.

* Vehicle Price Prediction using Visual Features (Yan et al., 20182)
  + The study makes the following contributions:

1. **New Datasets**: The first contribution is the creation of two original price-labeled datasets of 21,843 bicycles and 1,400 cars, specifically designed for visual price prediction.
2. **Effective Deep Learning Models**: The second contribution is the authors' custom PriceNet architecture, demonstrating that CNNs are highly effective at predicting prices from product images.
3. **Visual Explanations**: The third contribution involves revealing the specific image regions that the models focus on when determining price, helping us understand which visual features influence the model's price prediction.

**3.3 Identify potential challenges, best practices in the project**

1. **Data Collection**

**4.1 Identify and gather datasets containing second-hand guitar images and corresponding price data**

* Choose the Guitar Center website as the data source, which offers 4000+ guitar images by featuring reasonably priced items rather than luxury options.
* Choose Acoustic Guitar for model training, as Electric Guitar prices are often determined by intricate electric parts not visible in images.

**4.2 Gather second-hand guitar images from sources**

* Automatically download the mhtml files of Guitar Center web pages to the local.
* Retrieve the links of guitar images from local web pages.
* Download 4000+ second-hand guitar images using the links.

1. **Data Preprocessing**

**5.1 Automatically remove the background of guitar images**

**5.2 Manually eliminate invalid images (e.g., electric guitars, dummy images)**

**5.3 Crop the images to keep the top soundboard of guitars exclusively**

* Choose to focus our model training on the top soundboard of the guitar, as suggested by a guitar instructor, given that the fingerboard has minimal impact on price.
* Design an algorithm for batch cropping guitar images according to the above requirements.

**5.4 Resize and pad the guitar images**

* Consider identifying the largest image and add white padding to the other images to match its dimensions.
* Choose to downsample all images to 200 pixels in height, then add white padding to reach a uniform width of 200 pixels.

1. **Model Optimization and Fine tuning**

**6.1 Develop baseline deep learning models for prediction**

* Create a two-layer fully connected neural network as a toy model to verify preprocessing steps and dataset splits.
* Implement ResNet-50 as the baseline model to set a performance threshold for future improvements (Lecture slide W6D2-CNN.pdf, Comparing Complexity).

**6.2 Implement techniques to optimize model performance**

**6.3 Explore automated methods for feature selection and extraction to enhance model accuracy**

1. **Model Evaluation and Testing**

**7.1 Evaluate model performance using appropriate metrics, such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE)**

**7.2 Conduct cross-validation and validation testing to assess model generalization**

**7.3 Perform A/B testing or comparison against baseline methodologies to gauge improvement**

1. **Interpretation and Analysis**

**8.1 Interpret model predictions and analyze factors influencing price prediction accuracy**

**8.2 Investigate the impact of instrument-related features on model performance**

**8.3 Identify areas for further improvement and potential avenues for future research**

1. **Documentation and Reporting**

**9.1 Document data collection and preprocessing procedures, model development, and evaluation methodologies**

**9.2 Prepare comprehensive reports summarizing research findings, methodologies, and results**

1. **Presentation**

**10.1 Develop presentations to communicate project objectives, methodologies, and results**

1. **Revision and Iteration**

**11.1 Incorporate feedback from project sponsor/Advisor and peer to refine methodologies and models**

**11.2 Iterate on the project plan and objectives based on insights gained during the research process**