1. **Загрузить данные с изображениями людей в опасных/безопасных зонах и со средствами защиты и без, загрузить данные в data loader.**
2. **Проверить данные на выбросы и аномалии и убедиться, что обе категории представлены в достаточных объёмах - люди в опасных/безопасных зонах, со средствами защиты и без.**
3. **Разбить данные на три выборки: обучающую, валидационную и тестовую. Валидационная выборка нужна для подбора оптимальных гиперпараметров.**
4. **Найти подходящие предобученные модели обнаружения объектов (Single Shot Multibox Detector / You Only Look Once).**
5. **Разработать модели классификации, использующие различные алгоритмы, такие как свертывающие сети, метод решающих деревьев, градиентный бустинг, метод опорных векторов и другие.**
6. **Подобрать оптимальные гиперпараметры c помощью GridSearch, Tree-structured Parzen Estimator, Random Search, Байесовской оптимизации и Population Based Training.**
7. **Использовать data augmentation, чтобы снизить вероятность переобучения и сделать алгоритм более устойчивым к изменившимся условиям, например другому расположению камер.**
8. **Поскольку в таких алгоритмах обычно человек перепроверяет алгоритм, модель должна обращать большее внимание на ложноположительные результаты (ошибки первого рода), чтобы не пропустить потенциально опасные ситуации.**
9. **Запустить модель на тестовой выборке. Оценить точность предсказаний с помощью ROC AUC кривой, чтобы избежать ошибок, вызванных неравными классами.**
10. **Сравнить эффективность различных моделей и выбрать модель, показывающую наилучшую точность в соответствии с указанными потребностями.**
11. **Приоритет при выборе модели следует отдавать алгоритму с возможностью продолжающегося обучения, чтобы модель использовала результат человеческой перепроверки как дополнительные данные и впоследствии улучшала результат.**