

# CDK Data Stream AI

## Fall Progress Report



Group 65

Jacob Geddings, Inhyuk Lee, Juan Mugica

# Week 1-2: Class Start

Voted on Projects:

Group Interest: Artificial Intelligence

CDK Data Stream AI

- Company interested in exploring AI
- Handling thousands of data inputs
- Predictive Analysis
- Anomaly Detection

Class Tools:

OneNote

-Personal

-Group

Latex

# Week 3-4: Group Assignment, Meeting the Client

Group 65 Information:

Internal communication: Group text chat

External communication: Student email

Physical meeting times:

Weekdays 8-11 am

CDK Global - Chris Smith:

Initial meeting, one hour, in person.

Schedule set:

Friday 10:30 am Webex recurring

Clarity on project

# Week 4: Problem Statement

Problems with the Problem Statement:

- We were wrong on the assumed project goals prior to client meeting
- Encountered problems with using Latex for the first time

What we got out of it:

- What our client wants to have completed
- That our project is research heavy

# Week 5-6: Requirements Document

The goal of this project:

- Signature verification: signed, not signed, multiple signers

Stretch goals to analyze:

- Licenses (Legal)

- Vehicles (Make, Model, Damaged)

# Week 6: Client Meeting Summary

Priorities for the team shifted:

- Our primary concern is signature detection
- Container software, user interface, and cloud hosting not important
- Concerns raised over the primary language choice of C++
  - Backup plans being made for Python
- Client had to travel for a few weeks, anticipate future travelling as well

# Week 7-8: Technology Review

What each of us did:

Geddings:

- OpenCV Language

- Image Filter/Blur

- File Conversion

Lee:

- Deep Library

- Text Recognition

- User Interface

Mugica:

- Algorithm

- Container Software

- Development Platform

# Week 9-10: Design document

What we gained from it:

- Complete grouping of individual technology pieces
- Perspective on how each piece fits in with the project
- Forecast of how each will apply in the future

What was dropped:

- After client meetings several pieces were dropped by request
  - Container software, user interface, and cloud storage



# Design Doc: Computer Vision Library

Computer Vision Library choices:

- DL4J (Java)

- Tensor Flow

- OpenCV

# Design Doc: Languages of OpenCV

OpenCV supports the following languages

- C++

- Python

- Java

# Design Doc: Image Filtering Algorithm

Bilateral filtering:

- The bilateral filtering algorithm is highly proficient at retaining edges within an image and blurring the leftover less-defined elements
- Other methods were discarded due to concerns of over-blurring input images

# Design Doc:Text Recognition Method

## Intelligent Character Recognition

### -Three Steps

- Separation by line breaks and paragraphs

- Splitting text into individual characters

- Comparing against known characters and converting

# Design Doc: Image Converting

Image Magick:

- Standalone conversion framework compatible with c++
- Open source

# Design Doc: Runtime Environment + Neural Net algorithm

OSU's flip server

- Available 24/7
- Maintenance performed by professional staff

Convolutional Neural Networks

- Efficient at recognizing qualities within images
- Fast but memory inefficient

# Progress on OpenCV

## Setup process:

- Finding distribution and transferring to putty
- Utilizing cmake to build from source
- Configuring X11 with Xming
- Cmake implementation and familiarization

## Programs

- Created image displaying program for group to test installation
- Created image pixel analysis program via the “sliding window” strategy

# Winter Break Plans

## Meetings:

- Client meetings will become bi-weekly
- Group meetings will take place weekly

## Project plans:

- Creation of windows over signature boxes
- File conversion for PDF to JPEG
- Potentially work on image filtering