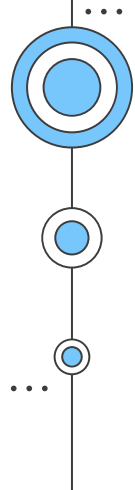# CloudFlex

How we helped to **raise 500k$** and **optimise infrastructure costs** from 15k to 200$/month **for US AI startup**
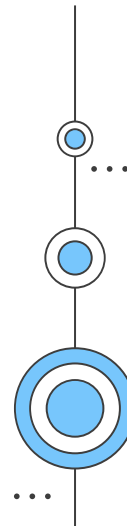
By CloudFlex.team. Team of Engineers who care about products and how they perform
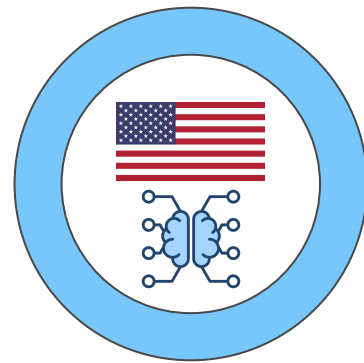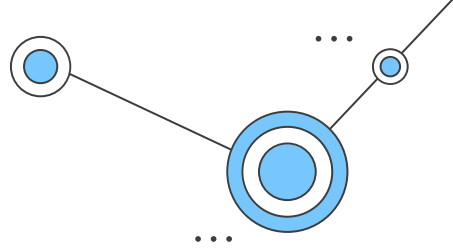
01
About client

Background

In the middle of 2022 we've started working with the **US AI startup** in the retail sector.

The request was to **provide ML expertise** and **help with cost reduction** on the backend as well as develop web part
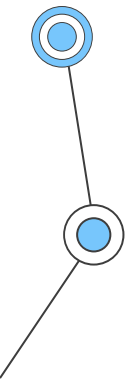
Who is the client?

ML startup from the US aggregating retail data from various sources leveraging advanced Machine Learning to predict style of clothing based on user preference.

A data collection pipeline consists of multiple stages responsible for transformations and enrichment. Given a huge volume of semi-structured data, some pipelines could take days for processing.

02
Challenges

Challenges Client Faced

Lack of AI expertise in-house

...

Reduce Infrastructure Costs

...

500k investments are blocked by urgent AI development

...

Business Goals

01 **AI part**

AI model development and optimization of feature extraction process

02 **Infrastructure optimization**

Infrastructure cost optimization (with pretty solid DevOps part)
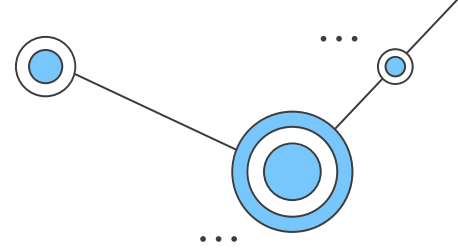
03 **New Features**

Satisfy urgent request and speed the things up

04 **Raise 500k$**

To raise investments it was required to finish AI work, urgent features and fix infrastructure during next 3-6 month
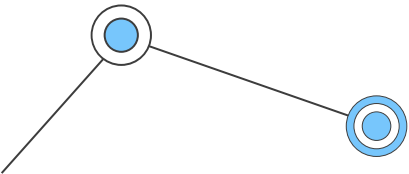
Few Words About Me

My name is **Kostiantyn - CloudFlex Co-Founder.**

With more than 8+ years of overall experience, **I am helping companies to build custom AI, Mobile, Web Solutions**.

At CloudFlex - we have strong knowledge of how to create custom intelligent systems or optimise their performance
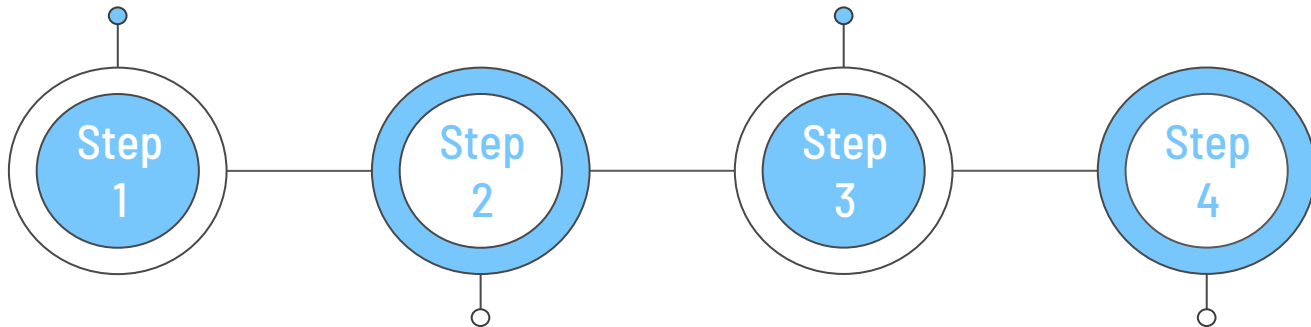
03

The approach

ML expertise

Started from the roots - requirements. This helped us to learn faster about how the AI core **should function**

Ultimate proposal on how to rebuild the system and still meet the deadlines during 3 month

Step 1

Step 2

Step 3

Step 4

Several sprints of learning and maintaining of existing codebase

Requested ML features were completed on time. **+500k$** raise during series round

Infrastructure 15k$ -> 200$ decrease

1. Design to support parallel computations

Design&Develop custom parallelized module that
splits compute-intensive operations into parts

2. Execute concurrently

Parallelize the process using AWS lambda – a small executor
capable of processing limited batches of data. Solution
reduced processing time x6 to 2 weeks

3. Harvest results

This **approach allowed for nearly 100% utilisation of
cluster capabilities** and extremely increased the overall
performance. Total computation cost has not exceeded
200$ comparing to 15k previously

04

The results

Results. AI part

Rewritten AI core in 3 month

9m+ of development with previous vendor was replaced by the solution we've built during ~3 month

Met all the deadlines
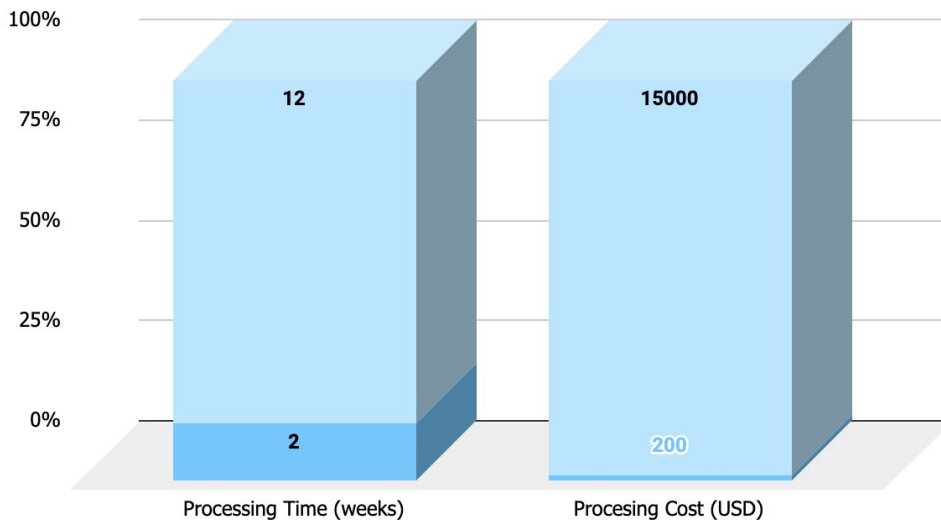
Precisely.

Company Raised 500k$
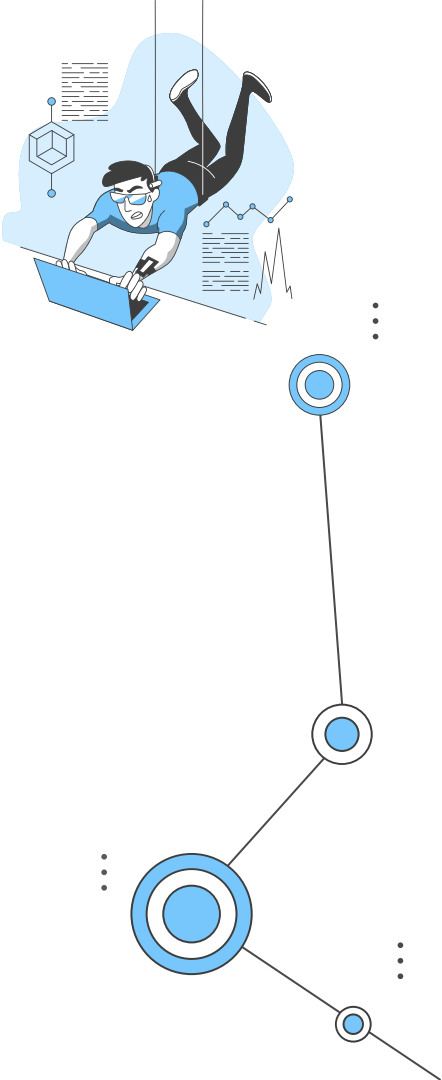
Immeasurable increase in trust

Results. Infrastructure

- 3 months turned into a 2 weeks of processing
- 15k$ showstopper on infrastructure became 200$ only

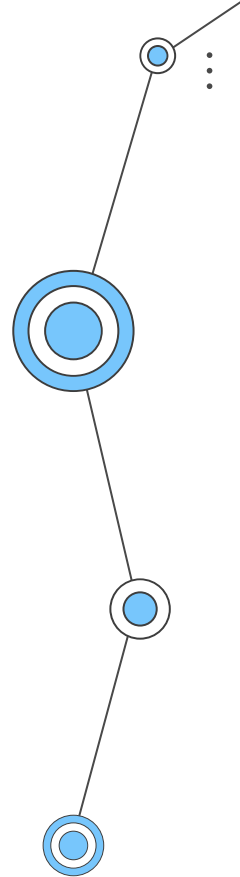**Evaluation of infrastructure optimization**

6x times faster

Processing pipeline

75x cheaper

15k VS 200$

Next Steps

Why don't we have a quick conversation so I could share with you common mistakes in project development that you could avoid?

I will also answer your questions that you might have about our expertise and how it may solve your business tasks.

BOOK CALL HERE: https://calendly.com/kostia-boskin/30min

Ready to start development right now?

Mention it, and we will make free analysis for you!