

Raportul tehnic al proiectului
“Machine learning using Cloud”
Adam Cristian A1
Pantiru Gabriel Alexandru A4

1. *Studiu de caz*
2. *Detalii tehnice*
3. *Tehnologii folosite*
4. *Business canvas*
5. *Diagrama Arhitecturala*
6. *Diagrama Use - Case*

1. Studiu de caz

Proiectul nostru constă în dezvoltarea unei aplicații user-friendly pentru utilizatorii care doresc să folosească servicii de clasificare date fără a fi nevoiți să aibă nevoie de cunoștințe tehnice de programare sau să știe foarte multe informații legate de machine learning.

Ideea proiectului a pornit de la o simplă discuție cum că de ce nu am putea folosi puterea mare de calcul a cloud-ului pentru a le putea oferi utilizatorilor o aplicație care să-i ajute să-și rezolve problemele cu ajutorul învățării automate. Domenii vizate ar fi cele de cercetare științifică, dar nu pot fi excluse și alte medii de lucru unde apar probleme de clasificare care ar conduce la ușurarea rezolvării unor taskuri.

La momentul actual cred ca am găsit o nișă, de pe urma careia am putea sa ajutam si pe alții folosind serviciile și api-urile furnizate de google cloud, care duc la mărirea performanței aplicației și la reducerea riscului de incertitudine cu privire la metodele de aplicare a algoritmilor de clasificare.

Mulți consideră machine learning ca fiind un subiect foarte complicat și greu de înțeles și aplicat, iar scopul aplicației noastre este de a crea un pod de legatura, în sensul în care machine learning ar putea devine ușor de folosit pentru oricine care are un dataset deja clasificat și pe baza lui să clasifice alte elemente asemănătoare.

2. Detalii tehnice

Cum am menționat și mai sus, aplicația va fi hostata pe cloud-ul celor de la google, folosind în mare parte serviciile si api-urile bazate pe machine learning ale lor (AutoML, AutoML Tables, Cloud Vision, Cloud Natural Language, Cloud Video Intelligence). Acestea permit ca aplicația să ruleze fără vreo intarziere mai ales la crearea unui model nou pe baza dataset-ului primit de la utilizator și asigura siguranța datelor utilizatorului, baza de date fiind tot pe cloud (Firestore).

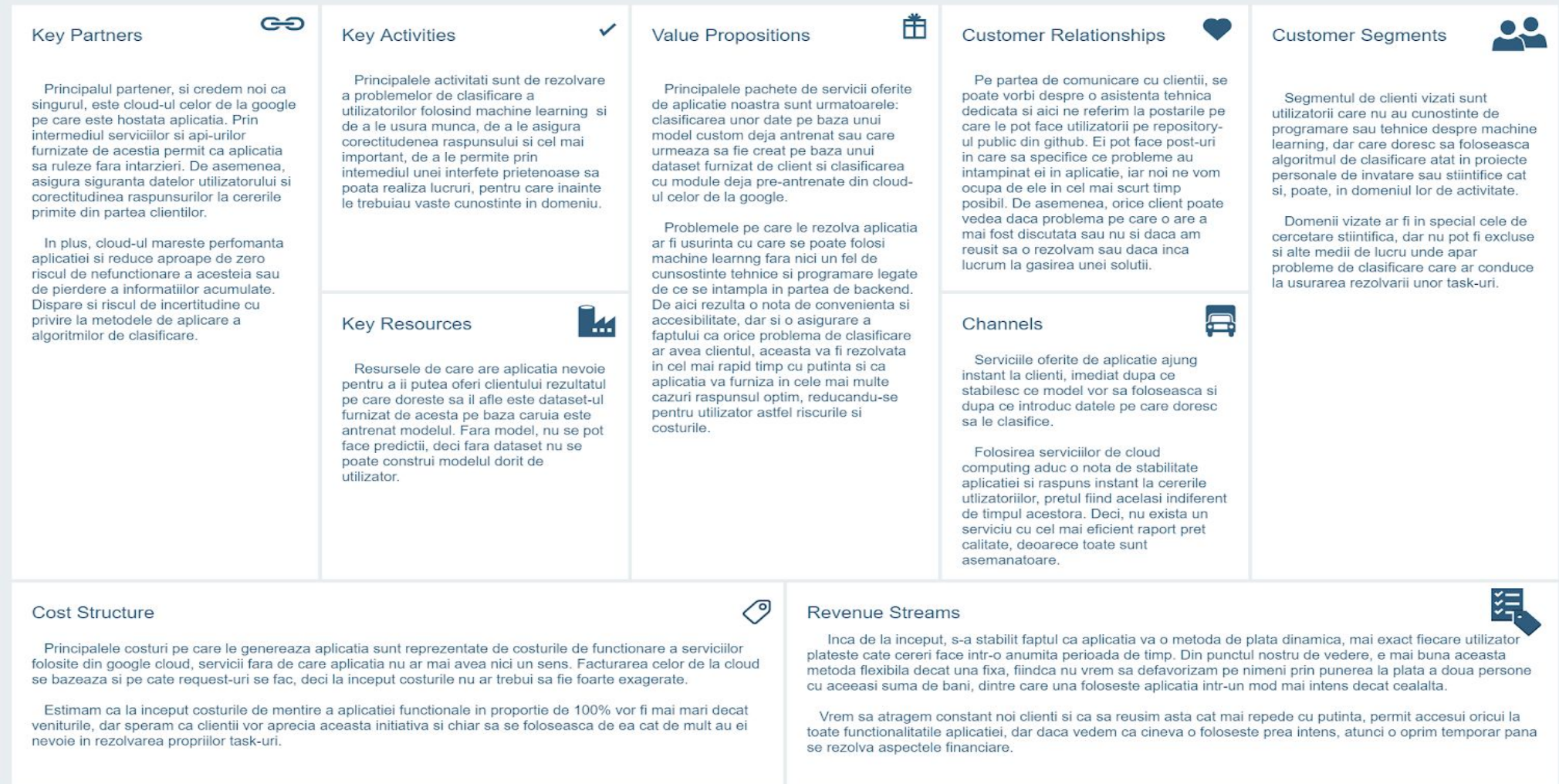
Principalele servicii furnizate de aplicatia noastra sunt clasificarea unor date pe baza unui model custom deja antrenat sau care urmează sa fie creat pe baza dataset-ului clientului și clasificarea cu module deja antrenate și furnizate direct de cloud. Pentru crearea unui model, utilizatorul furnizează un dataset de tip csv care apoi este folosit în procesul de antrenare. După ce modelul este creat folosind AutoML, utilizatorul poate sa ceara generarea unei predicții peste modelul deja creat. Pentru celălalt tip de clasificare, se vor folosi api-uri specifice precum Cloud Natural Language pentru recunoașterea entităților, Cloud Vision pentru determinare etichete imagini și Cloud Video Intelligence pentru determinare etichete video-uri.

3. Tehnologii folosite

- AutoML - antrenare model personalizate de învățare automata folosind-se expertiza necesara si intr-un timp cat mai scurt
- AutoML Tables - antrenare model învățare automată cu date structurate
- Cloud Vision API - api de detectare etichete imagine
- Cloud Natural Language API - api de analizare structura și semnificația textului pentru recunoașterea entităților
- Video AI - detecția etichetelor unui video
- Cloud Run - platforma pentru rulara si scalerea containerelor stateless
- Cloud Functions - funcții care răspund la evenimente din cloud
- Cloud Storage for Firebase - depozitare obiecte pentru stocarea și servirea conținutului generat de utilizator

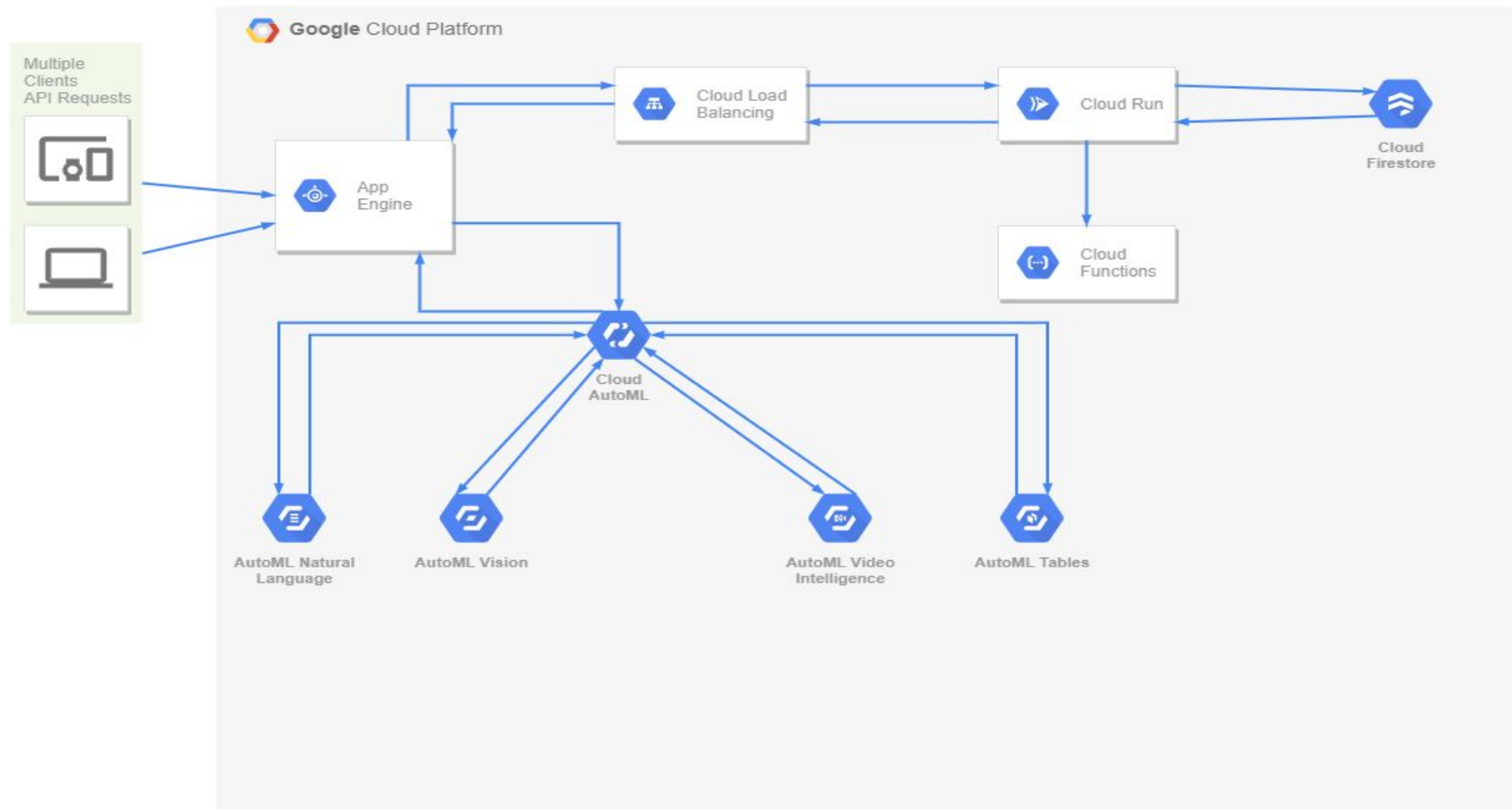
4. Business canvas

The Business Model Canvas



5. Diagrama arhitecturală

Architectural Diagram



6. Diagrama Use - Case

