**Machine Learning**

**Voortgang en taakverdeling**

**Woensdag 30 maart 2016**

**Vooruitgang afgelopen week:**

* Diede: featurevectoren van alle images zijn gemaakt.
* Daniëlle: SVM op basis van de gemiddelden van de featurevectoren is af. Submission gedaan: F1 score van 0.64327. Modelnaam: “DirecteRoute-SVM-gemiddelde”
* Laurens: 100 clusters getraind op basis van de train data.
* Roos: beetje beter geworden, maar nog niet helemaal.

**Vooruitgang deze woensdag:**

* Laurens: Code om de data in te delen in de 100 clusters is af. Moet nog draaien. Dit geeft een histogram per business.
* Daniëlle: SVM om de histogram per business te classificeren. Is af, maar moet nog draaien.
* Roos: ensemble code geschreven. Op dit moment bevat deze de modellen: “DirecteRoute-SVM-gemiddelde” en “frequentie-model”. Het ‘frequentie-model’ kijkt naar hoe vaak elke label voorkomt in de train data, en geeft deze kansen aan de labels. Op het laatste moment vandaag dit model gesubmit: F1 score van 0.7879!! Plek 70!
* Diede: Code SVM op basis van de min, Q1, Q2, Q3, max van de featurevectoren per business. Model: “DirecteRoute-SVM-minQ123max”. Moet nog afgemaakt worden en draaien.

**Taken komende week:**

**Laurens:** het “cluster+SVM” model draaien, testen en tweaken.

**Daniëlle:**

- SVM draaien met histogram data, voor het “cluster+SVM” model.

- model maken dat per foto de labels voorspelt op basis van de featurevectoren. Dit geeft dus een verzameling van voorspellingen per business. Model naam: “DirecteRoute-SVM-perFoto”

**Roos:** (afhankelijk van het werk van anderen)

-werken aan het samenvoegen van de voorspellingen per business, zodat er 1 voorspelling per business is. Model: “DirecteRoute-SVM-perFoto”

-draai ensemble methode voor alle modellen die af zijn.

**Diede:**

- draai het oude “colorfeature” model nog een keer, maar geef dan de kansen in plaats van de classificatie, zodat deze ook toegevoegd kan worden aan de ensemble.

- Maak model “DirecteRoute-SVM-minQ123max” af en draai.

**Plan volgende week:**

Woensdag zijn het “cluster+SVM” model en het “DirecteRoute-SVM-minQ123max” model af. We maken dan ook het “DirecteRoute-SVM-perFoto” model af. En werken aan de ensemble om alle modellen samen te voegen.

Maandag is onze laatste dag om eraan te werken. We gaan de ensemble verder tweaken. En onze laatste submissions doen.

Dinsdag 12 april deadline submission.

Verslag deadline: maandag 18 april.

Presentaties: donderdag 21 april.