# Accuracy Comparison for diff Methods & features:

WITH TEXT only, as Input, TARGET feature:**INDUSTRY**

* BERT Embeddings and BERT Classification: - **Acc - 0.54**
* BERT Embeddings (without other features)
  1. Random Forest - **Acc** score - **0.30**
  2. SVM - **Acc** score - **0.35**

TEXT and job\_count linkedin\_employee\_count company\_funding linkedin\_follower\_count rank company\_continent\_encoded company\_size\_encoded company\_type\_encoded revenue\_range\_encoded company\_country\_encoded company\_state\_encoded company\_city\_encoded as Input Features used. TARGET feature:**INDUSTRY**

* BERT Embeddings and other features: -

RANDOM Forest- **0.3 (acc) w/o MinMax Scaler**

RandomForestClassifier(n\_estimators=500) w MinMax Scaler: **0.3082**

Logistic Regression w MinMax Scaler: **0.4619**

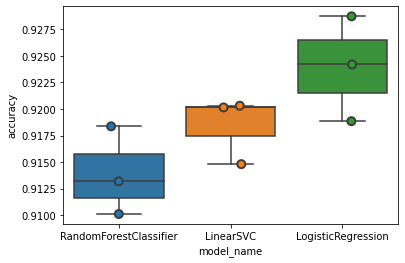
Linear SVM w MinMax Scaler: **0.4353**

XGBoost - **TLE**

On 147 classes:

|  | Accuracy(**BE**+other features) | Accuracy(**BE**+ **w/o** other features) | Accuracy (**tf-idf** + other features) |
| --- | --- | --- | --- |
| Random Forest | 0.31 | 0.30 | 0.32 |
| Logistic Regression | 0.46 | 0.46 | 0.28 |
| Linear SVC | 0.43 | 0.35 | 0.33 |

**RUN 1: 5 CLASSES**



**Food & Beverages 6270**

**Oil & Energy 5646**

**Apparel & Fashion 4692**

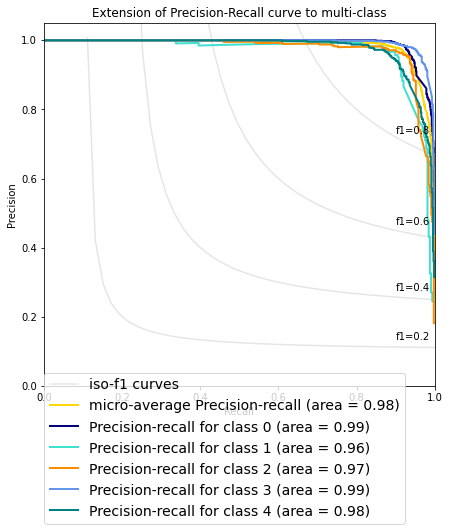
**Sports 3050**

**Music 2175**

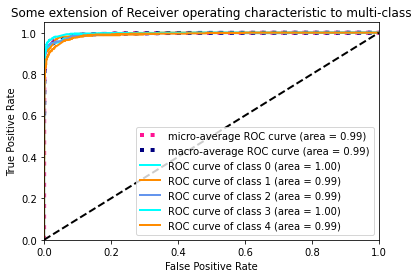
**Average precision score,**

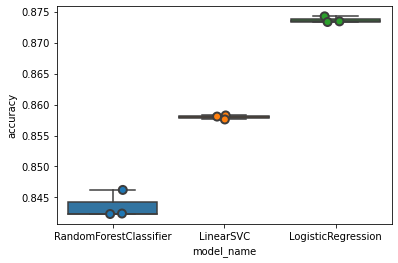
**micro-averaged over all classes:**

**AP=0.98**



**Precision Recall Curve ROC Curve**



**RUN 2: 10 CLASSES** 

**Food & Beverages 6270**

**Insurance 6230**

**Oil & Energy 5646**

**Apparel & Fashion 4692**

**Environmental Services 3271**

**Sports 3050**

**Mining & Metals 2797**

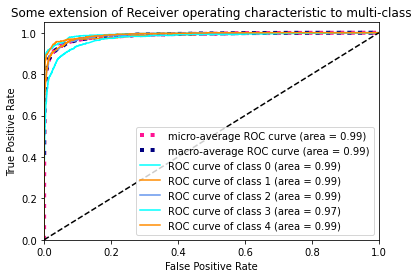
**Architecture & Planning 2480**

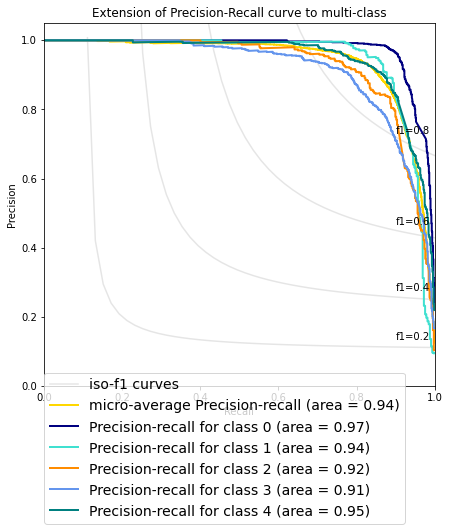
**Aviation & Aerospace 2402**

**Music 2175**

**'Average precision score, micro-averaged over all classes: AP=0.94'**

**Precision Recall curve ROC Curve**



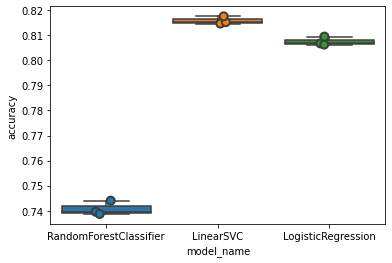


**RUN 3: 15 classes**

**Average precision score, micro-averaged over all classes: 0.91**

**Real Estate 10707**

**Hospital & Health Care 10023**

**Health Wellness and Fitness 8778**

**Medical Devices 6541**

**Food & Beverages 6270**

**Insurance 6230**

**Entertainment 5686**

**Oil & Energy 5646**

**Apparel & Fashion 4692**

**Environmental Services 3271**

**Sports 3050**

**Mining & Metals 2797**

**Architecture & Planning 2480**

**Aviation & Aerospace 2402**

**Music 2175**

**Average precision score, micro-averaged over all classes: 0.91**

