Implementation of a Portal Dedicated to Higgs Bosons for Experts and the General Public

DPG Spring Conference

Ivan Demchenko, Martin Kupka, André Sopczak, Antoine Vauterin,
Peter Žáčik

Czech Technical University in Prague

21 March 2022

Search for the Higgs boson

- 1989 – 2000: CERN - Large Electron-Positron Collider (LEP)
  - ▶ ALEPH
  - ▶ DELPHI
  - ▶ L3
  - ▶ OPAL
- 1987 – 2011: Fermilab – Tevatron accelerator
  - ▶ CDF
  - ▶ D0
- 2010 – present: CERN - Large Hadron Collider (LHC)
  - ▶ **ATLAS**
  - ▶ **CMS**
  - ▶ Ongoing Higgs measurements

Research

- 1000+ scientific publications (experimental results)
- Various types of experiments
- Various publishing methods
- New results each week
- Large number of articles - important to make a categorisation system

Goals of this project

1. Easy access to publications
2. Collection and **categorisation**
3. Visualisation of development precisions
4. Bringing the research closer to the public

Collecting data

What data?

- Publications – title, abstract, tables, graphs. . .
- Measured values – masses, productions modes, decay modes. . .

How?

- Fermilab (old websites) – web scraping
- CERN – CERN Document Server API
- Measured values – extract from text

Categorisation

- Goal of the publication
  - Experimental measurement
  - Search for "new physics"
- Observed events
  - Higgs boson production
  - Higgs boson decay
- Other properties
  - Number of collisions (luminosity)
  - Collision energy
  - Experiment
  - Current stage (preliminary, submitted, approved)

Natural Language Processing I

Vocabulary varies by the type of experiment

- Standard Model vs "new physics"
- Naive Bayes classifier
- Relies on the frequency of certain words
- No need to specify the words beforehand
- Training examples
- Tokenization, stopwords, lemmatization
- Python, scikit-learn, nltk

Natural Language Processing II

Numeric values and keywords detection

- Plain searching – ineffective
- Named Entity Recognition
- English corpus
- AI learns to recognize written text patterns
- Further algorithmic processing

NLP Example I

Search for charged Higgs bosons produced via vector boson fusion and decaying into a pair of W and Z bosons using proton-proton collisions at $\sqrt{s} = 13 \ TeV$.
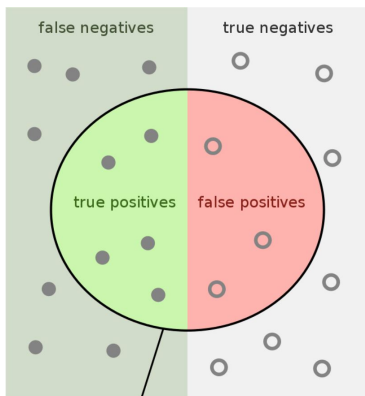
NLP Example II

Search for charged Higgs bosons produced via vector boson fusion and decaying into a pair of W and Z bosons using proton-proton collisions at $\sqrt{s} = 13\ TeV$.

Natural language processing III

- Categorisation model
- Manual training
- Training and testing set ~ 100 articles
- Standard F_1-score

Categorisation results I

1. $Precision = \frac{TP}{TP+FP}$
2. $Recall = \frac{TP}{TP+FN}$
3. $F_1 = 2 * \frac{P*R}{P+R}$

Categorisation results II

Category	Precision (%)	Recall (%)	F_1-score (%)
Luminosity	96	88	92
Energy	100	85	92
Production mode	87	85	86
Decay mode	81	79	80

Table: NER categorisation results

Web application

- Categorised publications stored in a database
- Daily updates
- API
- User Interface and administration

Web application

- Categorised publications stored in a database – **MongoDB**
- Daily updates – **Python cron jobs**
- API – **Flask**
- UI and administration – **React.js, Tailwind.css**

HBP – the portal for categorised information about the Higgs boson

The Higgs boson, named after the the physicist Peter Higgs is a subatomic particle. It is the manifestation of the Higgs field, a quantum field that is present everywhere throughout the Universe. Certain particles interact with the Higgs field via the Higgs mechanism and as a result obtain mass. The Higgs boson is the result of an excitation in this field.
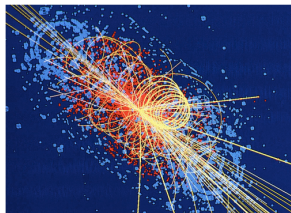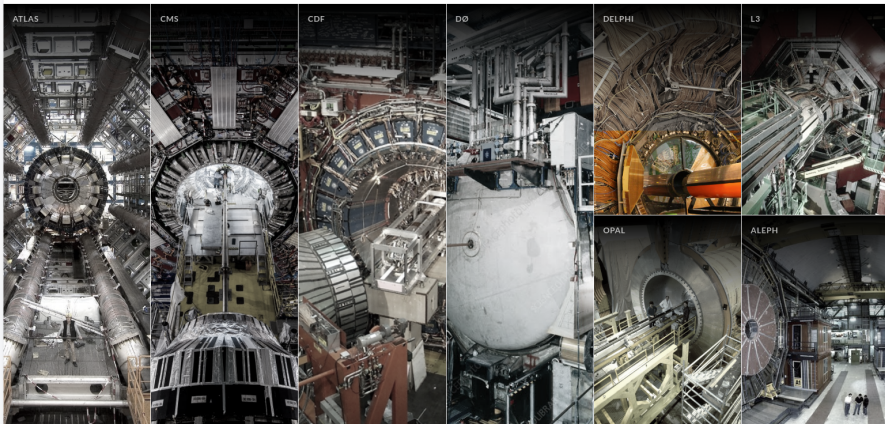
Source: CERN 2021

LEARN MORE

Illustration of a particle collision producing the Higgs boson
Lucas Taylor, CMS

ATLAS

CMS

CDF

DØ

DELPHI

L3

OPAL

ALEPH

Search

CMS **SUBMITTED** **BEYOND SM**

Search for charged Higgs bosons produced in vector boson fusion processes and decaying into vector boson pairs in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

A search for charged Higgs bosons produced in vector bosons and decaying into vector bosons, using proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ at the LHC, is reported. The data sa… See more

SAT APR 10 2021

PDF DETAILS

ATLAS **PRELIMINARY** **BEYOND SM**

Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector

Searches are performed for non-resonant and resonant di-Higgs boson production in the $b\bar{b}\gamma\gamma$ final state. The data set corresponds to an integrated luminosity of 139 fb^{-1} of … See more

WED MAR 31 2021

PDF DETAILS

ATLAS **PRELIMINARY** **SM**

Measurements of gluon fusion and vector-boson-fusion production of the Higgs boson in $H \rightarrow WW^{*} \rightarrow e\nu\mu\nu$ decays using pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector

The gluon fusion and vector-boson-fusion Higgs boson production modes are measured using proton–proton collisions in the $H \rightarrow WW^{*} \rightarrow e\nu\mu\nu$ decay channel. The Large Hadron C… See more

WED MAR 31 2021

PDF DETAILS

ATLAS **PRELIMINARY** **BEYOND SM**

Study of Higgs-boson production with large transverse momentum using the $H \rightarrow b\bar{b}$ decay with the ATLAS detector

This note reports studies of Higgs boson production with sizeable transverse momentum decaying to a $b\bar{b}$ quark pair. The analyzed data were recorded with the ATLAS detector in proton-proton coll… See more

SAT MAR 27 2021

PDF DETAILS

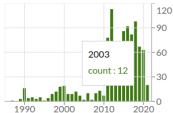**ATLAS** **PRELIMINARY** **BEYOND SM**

Search for Higgs boson decays into two spin-0 particles in the $b\bar{b}\mu\mu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13\,\text{TeV}$

This note presents a search for decays of the Higgs boson with a mass of 125 GeV to a pair of new spin-0 particles, $H \rightarrow aa$, … See more
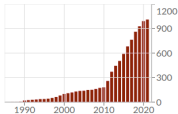
Deployment I
Components

What is required?

- Database hosting
- Web HTTP(S) server
- Website static files (HTML, CSS, Javascript)
- Server with enough computational power for categorisation (daily)

Deployment II
CERN PaaS Application Hosting service

- Components as Docker containers
- Container orchestration using **OpenShift**
- Automatic build from a GitHub repository
- One click deploy
- TLS CERN domain - secure HTTPS communication
- Categorisation runs on a shared CERN cluster
- Part of CPPP (Czech Particle Physics Project)

Deployment III

cern.ch/higgs

Thank you for your attention!