



OPENCLASSROOMS

Formation Ingénieur Machine Learning

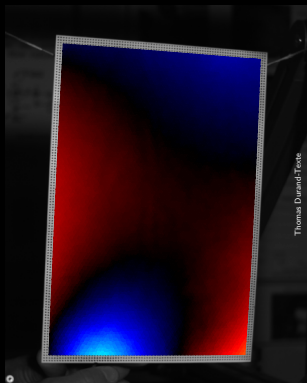
Projet: Définissez votre stratégie d'apprentissage

5 Janvier 2023

thomas.durandtexte@protonmail.com

Mon parcours

Parcours



Diplôme ingénieur

Master international

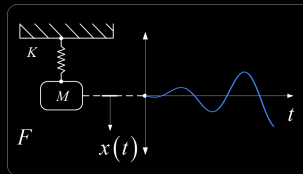
Thèse

4 ans recherche
post-doctorale



Parcours

- Ingénieur spécialisé Vibro-acoustique



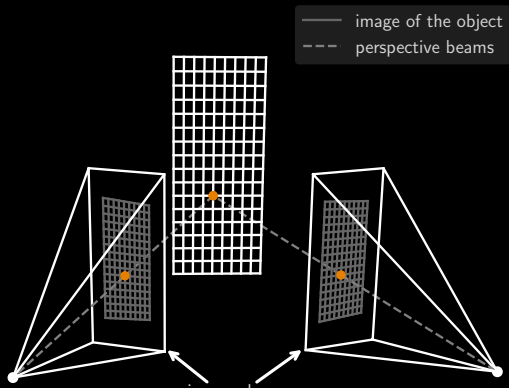
Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique



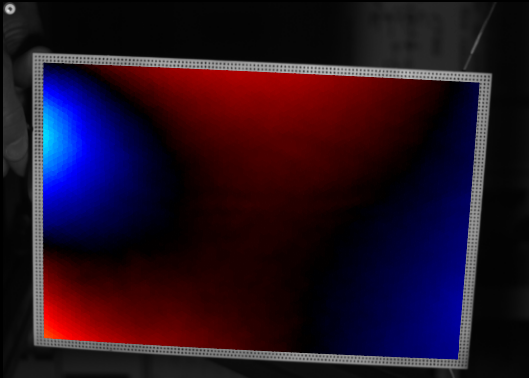
Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D



Parcours

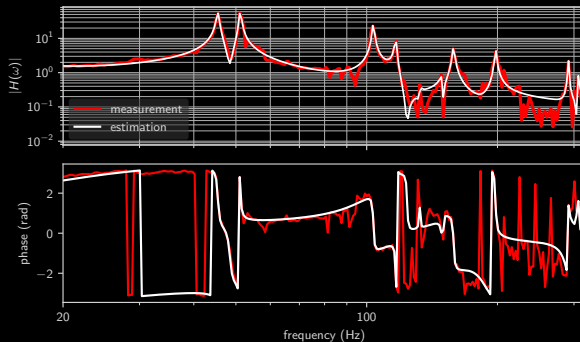
- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D



Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D

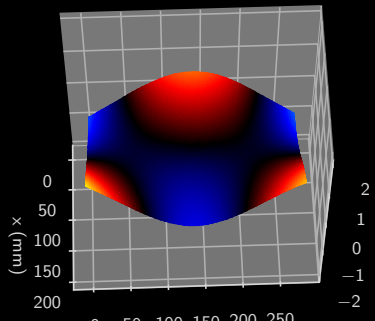
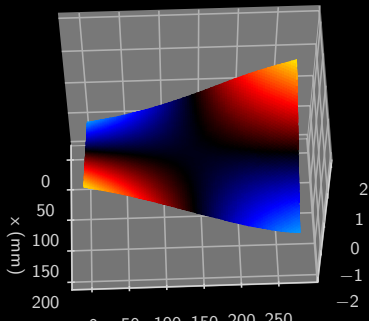
Estimation pour 1 point



Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D

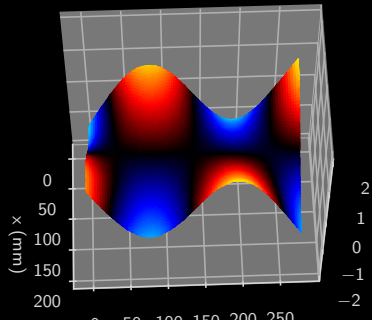
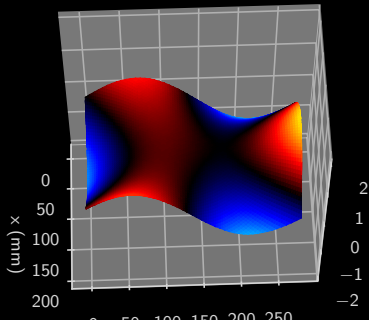
Estimation par mode



Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D

Estimation par mode



Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D
- Recherche post-doctoral (4 ans) :

Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D
- Recherche post-doctoral (4 ans) :
 - Imagerie infra-rouge appliqué à la dissipation vibratoire
 - Création de modèles : mécanique & thermique (et couplage des 2)
 - Étude de vibrations de céramiques

Parcours

- Ingénieur spécialisé Vibro-acoustique
- Master international en Électro-Acoustique
- Doctorat : mesure de vibrations par vision 3D
- Recherche post-doctoral (4 ans) :
 - Imagerie infra-rouge appliqué à la dissipation vibratoire
 - Création de modèles : mécanique & thermique (et couplage des 2)
 - Étude de vibrations de céramiques
- Recherche personnelle ?

Pourquoi une nouvelle formation ?

Une nouvelle formation pour :

- Le plaisir d'apprendre



Une nouvelle formation pour :

- Le plaisir d'apprendre
- Élargir mon champs de compétence



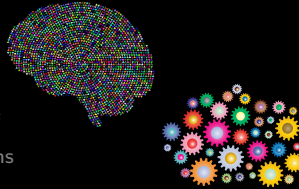
Une nouvelle formation pour :

- Le plaisir d'apprendre
- Élargir mon champs de compétence
- Pouvoir accéder à plus d'applications



Une nouvelle formation pour :

- Le plaisir d'apprendre
- Élargir mon champs de compétence
- Pouvoir accéder à plus d'applications



Objectif après la formation :

Créer une micro-entreprise (freelance)

Pourquoi OpenClassrooms ?

Une formation OpenClassrooms c'est :

- 45 minutes / semaine de mentorat avec un professionnel

Une formation OpenClassrooms c'est :

- 45 minutes / semaine de mentorat avec un professionnel
- En distanciel (souplesse)



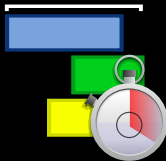
Une formation OpenClassrooms c'est :

- 45 minutes / semaine de mentorat avec un professionnel
- En distanciel (souplesse)
- Un accès à une communauté



Une formation OpenClassrooms c'est :

- 45 minutes / semaine de mentorat avec un professionnel
- En distanciel (souplesse)
- Un accès à une communauté
- Un apprentissage par **projets**



Les projets

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)
2. Anticipez les besoins en consommation de bâtiments
(Modèle d'apprentissage supervisé)

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)
2. Anticipez les besoins en consommation de bâtiments
(Modèle d'apprentissage supervisé)
3. Segmentez des clients d'un site e-commerce
(Modèle d'apprentissage non supervisé)

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)
2. Anticipez les besoins en consommation de bâtiments
(Modèle d'apprentissage supervisé)
3. Segmentez des clients d'un site e-commerce
(Modèle d'apprentissage non supervisé)
4. Catégorisez automatiquement des questions
(Gestion de données non structurées & extraction de features)

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)
2. Anticipez les besoins en consommation de bâtiments
(Modèle d'apprentissage supervisé)
3. Segmentez des clients d'un site e-commerce
(Modèle d'apprentissage non supervisé)
4. Catégorisez automatiquement des questions
(Gestion de données non structurées & extraction de features)
5. Classez des images à l'aide d'algorithmes de Deep Learning
(Gestion d'images & Deep Learning)

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)
2. Anticipez les besoins en consommation de bâtiments
(Modèle d'apprentissage supervisé)
3. Segmentez des clients d'un site e-commerce
(Modèle d'apprentissage non supervisé)
4. Catégorisez automatiquement des questions
(Gestion de données non structurées & extraction de features)
5. Classez des images à l'aide d'algorithmes de Deep Learning
(Gestion d'images & Deep Learning)
6. Développez une preuve de concept
(Veille technologique & preuve de concept)

Les projets :

1. Concevez une application au service de la santé publique
(Gestion de données incomplètes)
2. Anticipez les besoins en consommation de bâtiments
(Modèle d'apprentissage supervisé)
3. Segmentez des clients d'un site e-commerce
(Modèle d'apprentissage non supervisé)
4. Catégorisez automatiquement des questions
(Gestion de données non structurées & extraction de features)
5. Classez des images à l'aide d'algorithmes de Deep Learning
(Gestion d'images & Deep Learning)
6. Développez une preuve de concept
(Veille technologique & preuve de concept)
7. Participez à une compétition Kaggle !
(Mise en pratique, gestion & communication de projet de développement)

Les projets les plus **motivants** :

Concevez une application au service de
la santé publique



alimentation, prise en main des
outils

Les projets les plus **motivants** :

Concevez une application au service de
la santé publique



alimentation, prise en main des
outils

Anticipez les besoins en consommation
de bâtiments



modélisation, énergie

Les projets les plus **motivants** :

Concevez une application au service de
la santé publique



alimentation, prise en main des
outils

Anticipez les besoins en consommation
de bâtiments



modélisation, énergie

Classez des images à l'aide
d'algorithmes de Deep Learning



traitement d'images, (Deep
learning?)

Les projets les plus **difficiles** à priori :

Segmentez des clients d'un site
e-commerce



plus éloigné de mes
compétences

Les projets les plus **difficiles** à priori :

Segmentez des clients d'un site
e-commerce



plus éloigné de mes
compétences

Catégorisez automatiquement des
questions



plus éloigné de mes
compétences

Les projets les plus **difficiles** à priori :

Segmentez des clients d'un site
e-commerce



plus éloigné de mes
compétences

Catégorisez automatiquement des
questions



plus éloigné de mes
compétences

Participez à une compétition Kaggle !



compétition

Dates prévisionnelles de soutenances :

Import des packages et chargement des données

Les données correspondent à un temps estimatif de travail des différents projets

```
import pandas as pd
import numpy as np

data = pd.read_csv('projets.csv') ;
display( data )
```

Python

	projet	durée	durée cours	durée supplémentaire
0	Application santé	70h	25h	30h
1	Besoin conso batiements	100h	47h	40h
2	Segmentez clients e-commerce	70h	25h	20h
3	Catégorisez questions	80h	8h	80h
4	Classez images algo Deep Learning	100h	40h	60h
5	Développez une preuve de concept	30h	1h	20h
6	Compétition Kaggle	50h	0h	50h

Dates prévisionnelles de soutenances :

Vérification du type des colonnes et conversion des temps/durée str->int

- Toutes les colonnes sont vérifiées avec une boucle for
- si la colonne est déjà en int, ou si 'durée' n'est pas présent dans la key, alors la colonne est inchangée

Python

```
for key in data.keys():
    if (data[key].dtype == 'int64') or ( not 'durée' in key): continue ;
    data[key] = pd.to_numeric(data[key].str.replace('h',''), errors='coerce' ) ;
# for #
print('dtypes:\n', data.dtypes)
```

Python

```
dtypes:
projet                object
durée                 int64
durée cours           int64
durée supplémentaire  int64
dtype: object
```

Dates prévisionnelles de soutenances :

Calcul du temps total pour chaque projet

```
data['durée totale'] = data['durée'] + data['durée cours'] + data['durée supplémentaire'] ;  
display( data )
```

Python

	projet	durée	durée cours	durée supplémentaire	durée totale
0	Application santé	70	25	30	125
1	Besoin conso batiments	100	47	40	187
2	Segmentez clients e-commerce	70	25	20	115
3	Catégorisez questions	80	8	80	168
4	Classez images algo Deep Learning	100	40	60	200
5	Développez une preuve de concept	30	1	20	51
6	Compétition Kaggle	50	0	50	100

Dates prévisionnelles de soutenances :

Temps initial estimatif en heures

1. calcul du nombre de jours de janvier à juin inclu
2. calcul du nombre de semaines (entières)
3. calcul du nombre d'heures à disposition sur une base de 35h/semaine
4. somme des heures des différents projets (estimation) et comparaison avec le nombre d'heures à disposition

```
n_jours = np.sum( [31,28,31,30,31,30] ) ;
n_sem = n_jours // 7 +2 ; # floor + ajout 2 semaines
n_jours_restant = n_jours - 7*n_sem ;
n_heures_dispo = 35 * n_sem ;
heures_total_estim = data['durée totale'].sum() ;
print( 'nombre de jours initial : {}'.format(n_jours) ) ;
print( 'nombre de semaines : {}'.format(n_sem) ) ;
print( 'nombre de jours restants : {}'.format(n_jours_restant) ) ;
print( "nombre d'heures disponibles : {}".format(n_heures_dispo) ) ;
print( "nombre d'heures total estimé : {}".format(heures_total_estim) ) ;
```

Python

```
nombre de jours initial : 181
nombre de semaines : 27
nombre de jours restants : -8
nombre d'heures disponibles : 945
nombre d'heures total estimé : 946
```

Dates prévisionnelles de soutenances :

Estimation de la durée de chaque projet

Valeurs en semaine, sur une base de 35h/semaine, arrondi à l'entier supérieur

```
key = 'durée estimée (semaines)' ;  
data[key] = ( data['durée totale'] / 35 + 1. ).astype('int') ;  
print('nombre de semaines initial (janvier à juin) : {}'.format( n_sem ) ) ;  
print('nombre de semaines estimé : {}'.format( data[key].sum() ) ) ;  
display( data )
```

 Python Python

nombre de semaines initial (janvier à juin) : 27

nombre de semaines estimé : 30

	projet	durée	durée cours	durée supplémentaire	durée totale	durée estimée (semaines)
0	Application santé	70	25	30	125	4
1	Besoin conso batiements	100	47	40	187	6
2	Segmentez clients e-commerce	70	25	20	115	4
3	Catégorisez questions	80	8	80	168	5
4	Classez images algo Deep Learning	100	40	60	200	6
5	Développez une preuve de concept	30	1	20	51	2
6	Compétition Kaggle	50	0	50	100	3

Dates prévisionnelles de soutenances :

Estimation des dates de soutenance

1. Conversion des durées cumulées en *timedelta*
2. Ajout à une date initiale le 02/01/2023

```
duree_cumsum = 7*data['durée estimée (semaines)'].cumsum() ;  
data['date soutenance (estimation)'] = \  
|     pd.to_datetime( '02/01/2023',format='%d/%m/%Y') \  
|     + pd.to_timedelta( duree_cumsum.astype(str) + ' days') ;  
display(data)
```

Python

	projet	durée	durée cours	durée supplémentaire	durée totale	durée estimée (semaines)	date soutenance (estimation)
0	Application santé	70	25	30	125	4	2023-01-30
1	Besoin conso batiements	100	47	40	187	6	2023-03-13
2	Segmentez clients e-commerce	70	25	20	115	4	2023-04-10
3	Catégorisez questions	80	8	80	168	5	2023-05-15
4	Classez images algo Deep Learning	100	40	60	200	6	2023-06-26
5	Développez une preuve de concept	30	1	20	51	2	2023-07-10
6	Compétition Kaggle	50	0	50	100	3	2023-07-31

Dates prévisionnelles de soutenances :

Création de la chaine de caractère à copier/coller dans la fiche de début de formation

```
str_out = '' ;  
for i, date in enumerate( data['date soutenance (estimation)'].tolist() ) :  
    date_str = date.strftime('%d/%m/%Y') ;  
    str_out += '\n-Projet {:}: {}'.format( i+1, date_str ) ;  
# for i, date #  
print(str_out) ;
```

Python

```
-Projet 1: 30/01/2023  
-Projet 2: 13/03/2023  
-Projet 3: 10/04/2023  
-Projet 4: 15/05/2023  
-Projet 5: 26/06/2023  
-Projet 6: 10/07/2023  
-Projet 7: 31/07/2023
```

Quelques bonnes pratiques

- Utiliser un calendrier et noter les date butoires de fin de projets - Utiliser la technique POMODORO (pour limiter la fatigue) - Bien préparer les sessions de mentorat, notamment en tenant à jour une liste de questions - Ne pas rester bloqué sur un point du projet en cours, avancer sur d'autres points non bloquants en attendant la prochaine session de mantorat / poser des questions à d'autres étudiants via la plateforme SLACK

Outils collaboratifs

- GitHub - Mails - Plateform OPENCLASSROOMS - Discord

Thomas-Durand-Texte / OC-formation-Ingenieur-Machine-Learning Public

[Pin](#) [Unwatch 1](#) [Fork](#) [Star 0](#)

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

[main](#) [1 branch](#) [0 tags](#) [Go to file](#) [Add file](#) [Code](#)

Thomas-Durand-Texte upload

6a42a89 yesterday 3 commits

Projets	upload	yesterday
.gitignore	first commit	yesterday
README.md	first commit	yesterday
git_init	first commit	yesterday
git_upload	first commit	yesterday

README.md

Dossier partagé pour la formation OpenClassRooms Ingénieur Machine Learning

- Un dossier regroupe les différents fichiers de chaque projet (dans des sous-dossiers)
- Un dossier correspond aux informations partagées / autre ?

About

No description, website, or topics provided.

[Readme](#)
0 stars
1 watching
0 forks

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Languages

Jupyter Notebook 69.4%

TeX 29.7%

Python 0.9%

Merci pour votre attention.

Artworks

commons.wikimedia :

- Mass Spring System Resonance.gif by Guillermo Bossio

openclipart.org :

- loudspeaker by forestgreen
- Prismatic Cloud Gears 2, Abstract Brain Triangles Prismatic by GDJ
- Time Project by lbear
- Laptop Computer Icon by kael_179