

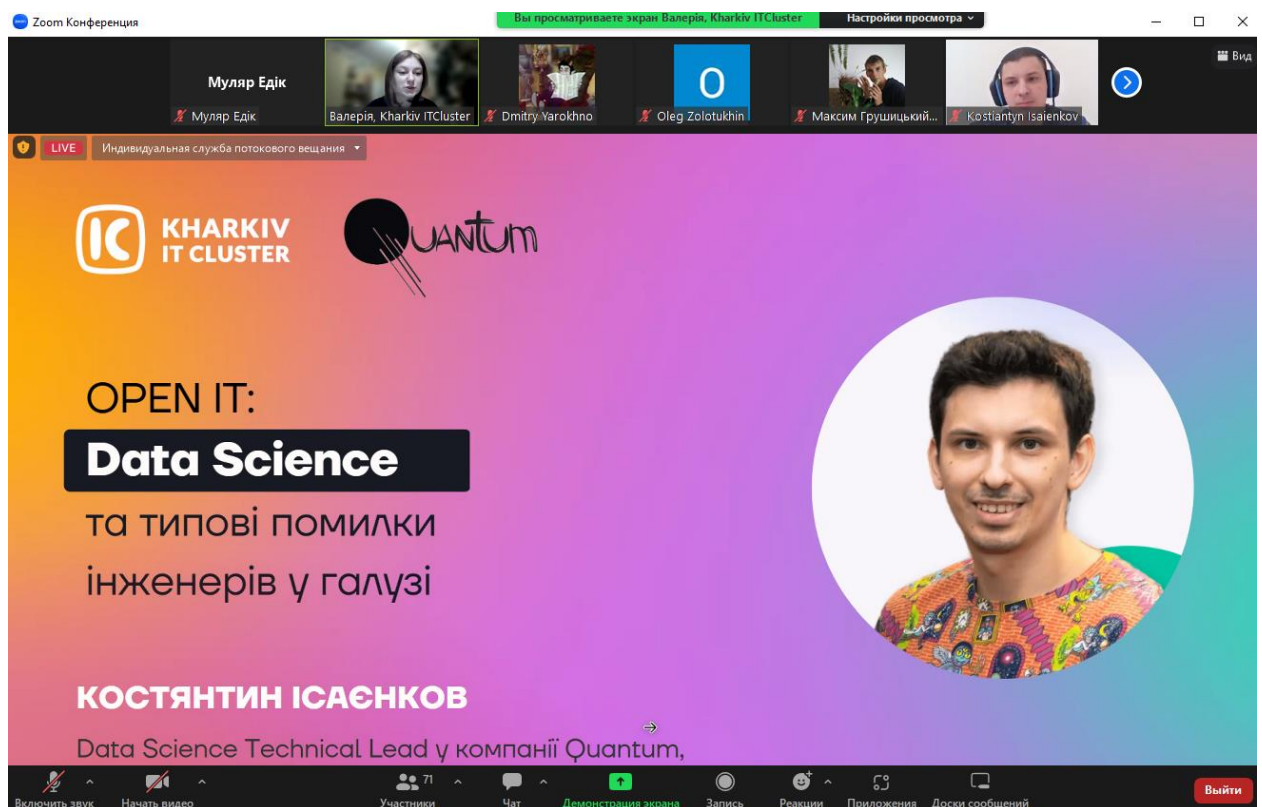
👉 IT Student — будь в курсі IT-можливостей!

📍 Викладачі та студенти ОП «Комп'ютерні науки» 15 лютого відвідали перший у 2023 році online-вебінар Kharkiv IT Cluster на платформі ZOOM з теми: «Data Science та типові помилки інженерів у галузі».

🗣️ Костянтин Ісаєнков — Data Science Technical Lead у компанії Quantum, засновник Data Science Factory та великий фанат Artificial Intelligence, розповів на зустрічі про:

- ? що таке Data Science?
- ? які типи помилок є у проектах Data Science?
- ? як почати кар'єру в Data Science?

Дякуємо за цікавий online-вебінар та вичерпні відповіді на питання!



Найактивніші учасники online-вебінару поділилися своїми враженнями:

👤 Едуард МУЛЯР, студент групи КНм-22-1

Сьогодні мною було відвідано вебінар, який проводив Kharkiv IT Cluster з теми: «Data Science та типові помилки інженерів у галузі». Даний вебінар проводив співробітник компанії Quantum: Data Science Technical Lead Костянтин Ісаєнков.

Основною тематикою вебінару був Data Science. Ментор Костянтин розповів що собою представляє даний напрямок в IT галузі, його перспективи та необхідність у сучасному IT. Найбільше мені сподобалося те, що ментор пояснював про Data Science на основі проектів, які розробляв він і його команда. Велику увагу було

приділено помилкам та необхідним вмінням в Data Science. Також гарно було описано спектр технологій, які необхідні при роботі в даному напрямку.

Фінальною частиною MeetUp були відповіді на запитання. Ментор Костянтин гарно та конструктивно відповідав на задані питання.

Для себе взяв до уваги важливість уміння працювати у команді, також обов'язково необхідно розвивати комунікативні навички, не боятися вивчати нові технології і завжди вірити в себе.

За вебінар ставлю 4.5/5.

🗨 Ольга ЗАЛУЦЬКА, студентка групи КНм-22-1

Чудовий захід від колег по цеху. Дякуємо, що в непростий час знайшли час для цікавої зустрічі, адже тема досить актуальна як і для студентів, так і для тих, хто вже працює в ІТ.

🗨 Богдан ВОНСОВИЧ, студент групи КН-20-2

Мені дуже сподобався online-вебінар проєкту Open IT, від Kharkiv IT Cluster. Я ознайомився з поняттям Data Science, типами помилок у проєктах Data Science і вимогами до початку кар'єри в Data Science. Також спікер давав вичерпні відповіді на запитання, які ставили слухачі online-вебінару.

Сподобалося, що online-вебінар був добре організований, було дуже зручно те, що він відбувався онлайн.

🗨 Максим ЗАПОРОЖЕЦЬ, студент групи КНм-22-1

Мені вебінар сподобався. Костянтин Ісаєнков дуже добре розказав основну інформацію про Data Science, було розглянуто що це та його особистий досвід при роботі над різними проєктами зі застосуванням цього. Також увагу приділили моменту працевлаштування, які потрібні знання для цього та навички. Вебінаром залишився задоволений.

🗨 Анастасія ЗАГОРОДНЯ, студентка групи КН-22-1

Була на вебінарі. Дуже цікава інформація, все чітко та лаконічно. Думаю це буде корисно у майбутньому.

🗨 Кіріл СПІВАК, студент групи КН-22-1

Відвідав з зацікавленістю вебінар, він мені сподобався тим, що спікер доступно пояснив про типові помилки не тільки у проєктах, але і під час самого підходу до вирішення задачі. Тобто, будь яка задача Data Science має починатися з візуалізації, а саме з аналізу даних які ми маємо. В двох словах, ставлю оцінку 5/5.

🗨 Богдан ДЕНИСЕНКО, студент групи КН-22-2

Захід дуже сподобався, оскільки сам цікавлюся даною сферою. Цікаво було послухати людину, яка вже працює в цій сфері. У структурованій презентації були висвітлені проблеми, типові помилки, подано актуальну інформацію про професію. Доповідач дав відгуки на всі запитання, що виникали. Залюбки відвідав би ще подібні заходи.

🗨 Юлія СОБКОВА, студентка групи КНм-22-1

Легко і доступно пояснили перші етапи влаштування на ІТ- спеціальність, та розповіли про свою компанію для студентів.

🗨 Ілля СКУЛЬСЬКИЙ, студент групи КНм-22-1

По суті цікава, але дуже узагальнена тема. Хотілось би побачити чи то приклад роботи(структура роботи) чи метод моделей розробки).

🗨 Вадим ОХОТА, студент групи КНм-22-1

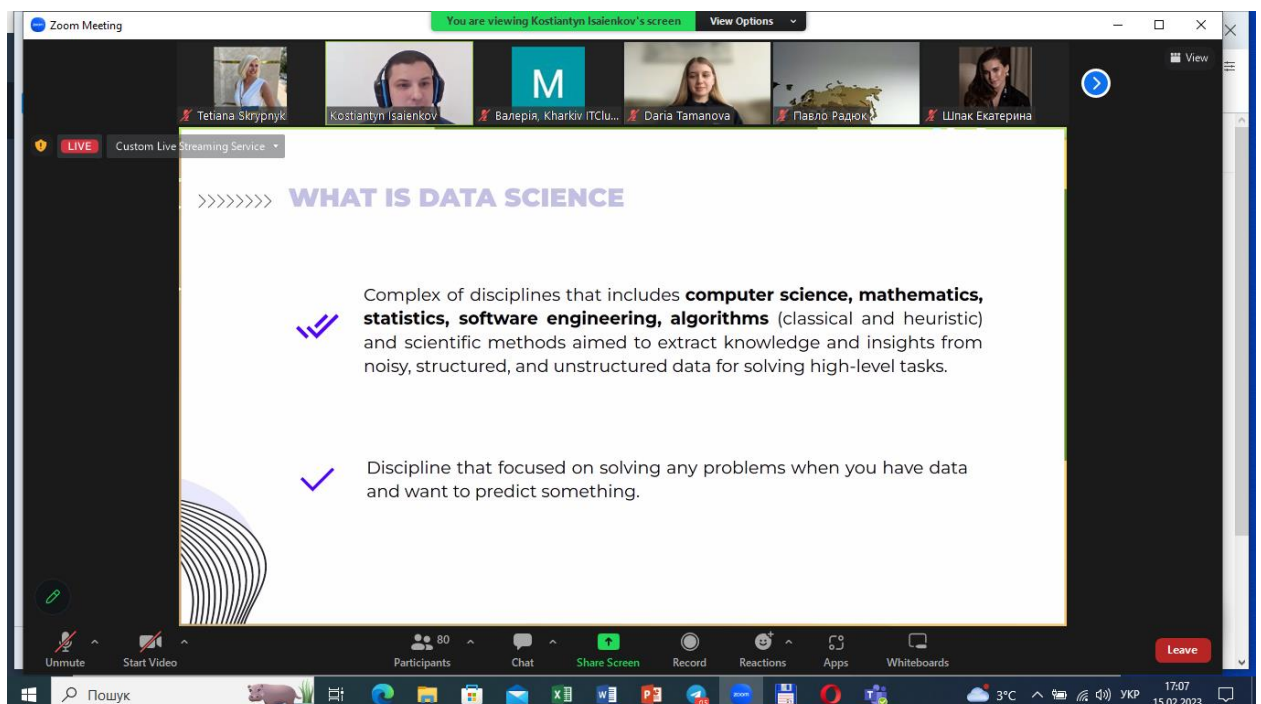
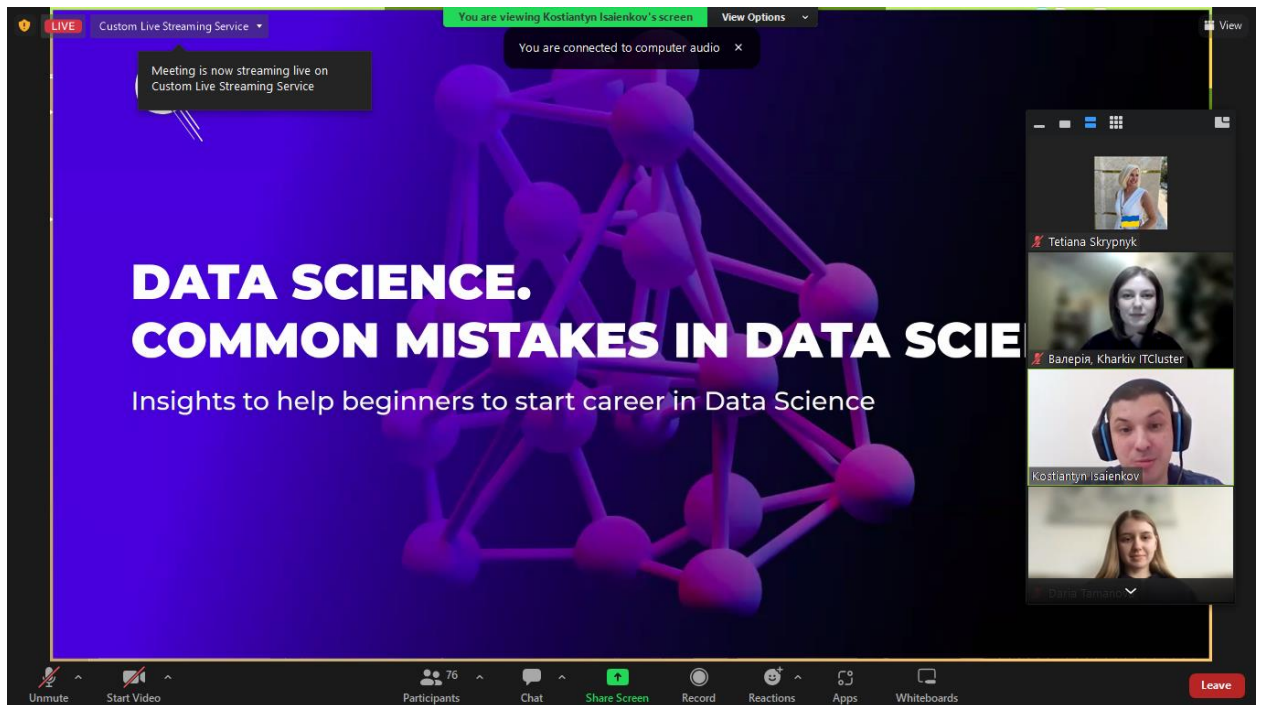
Ну в цілому задоволений вебінаром. Доступно подана інформація про роботу з датою. Чітко вказані нюанси та часті помилки в логіці з працею з ДБ. Для себе, як програміста, дізнався корисну інформацію яку можна використати в роботі

🗨 Володимир ДОБРЖАНСЬКИЙ, студент групи КНм-22-1

достатньо цікаво, хороша презентація, хороший екскурс в професію, наведені кейси тісно переплітаються і розробкою ПЗ, такі як відсутність юніт тестів

📎 **Володимир ДОБРЖАНСЬКИЙ**, студент групи КНм-22-1

достатньо цікаво, хороша презентація, хороший екскурс в професію, наведені кейси тісно переплітаються і розробкою ПЗ, такі як відсутність юніт тестів



Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Валерія, Kharkiv ІТClu... | Павло Радюк | Daria Tamanova | Шпак Катерина

LIVE Custom Live Streaming Service

DATA SCIENCE SPECIALIZATIONS

- Classical Machine Learning**
Direction focused on solving predictive and analytical problems on the tabular and similar data.
- Computer vision**
Tasks related to visual information like image processing, video analytics, etc.
- Natural Language processing**
Solving tasks linked with textual information.
- MLOps**
Solving tasks that should simplify processes of deploying ML solutions into production.

Unmute | Start Video | Participants (82) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:08 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Валерія, Kharkiv ІТClu... | Павло Радюк | Daria Tamanova | Шпак Катерина

LIVE Custom Live Streaming Service

COMPUTER VISION. DEFORESTATION

- #1**
Working with Sentinel-2 satellite imagery
- #2**
Use several years of data for training
- #3**
Cooperation with GEO experts
- #4**
Training Deep Learning models based on UNET architecture
- #5**
As a result – published scientific paper and dozens of potential customers interested in this topic

Unmute | Start Video | Participants (86) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:10 15.02.2023

Zoom Meeting You are viewing Kostiantyn Isaienkov's screen View Options

Tetiana Skrypnyk Kostiantyn Isaienkov Валерія, Kharkiv ІТClu... Павло Радюк Daria Tamanova Шпак Катерина

COMPUTER VISION. GREENHOUSES AUTOMATION

- #1 Prize winners in International Greenhouses simulation competition (Netherlands).
- #2 The algorithm is based on the combination of reinforcement learning, computer vision and domain expert knowledge.
- #3 Application inside the real environment with various number of sensors.
- #4 Currently, we focused on the productization of the solution.

Unmute Start Video Participants 87 Chat Share Screen Record Reactions Apps Whiteboards Leave

Пошук 2°C 17:12 15.02.2023

Zoom Meeting You are viewing Kostiantyn Isaienkov's screen View Options

Tetiana Skrypnyk Kostiantyn Isaienkov Валерія, Kharkiv ІТClu... Павло Радюк Daria Tamanova Шпак Катерина

CLASSICAL MACHINE LEARNING. AUTOML

- #1 The goal is to build a system that is able to solve classification, regression, clustering and time series forecasting problems on the tabular data
- #2 System should support full stack ML solution - starting from the data cleaning, feature engineering, model training and hyperparameters tuning
- #3 Smart feedback mechanism for dataset extension
- #4 Minimum interaction with user

Unmute Start Video Participants 88 Chat Share Screen Record Reactions Apps Whiteboards Leave

Пошук 2°C 17:15 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Валерія, Kharkiv ITClu... | Павло Радюк | Daria Tamanova | Шпак Катерина

NLP. GENE-DISEASE RELATION EXTRACTION

#1
The goal is to build a system that should identify the relationship between mutation of the gene and disease from a large medical text corpus

#2
Two-step sequential system:
1. Bert model for Named Entity Recognition (Gene, Disease, Mutation entities).
2. Bert model for Relation classification. Four types of relations (Neutral, Unknown, Provocation, Decreasing)

#3
Manual labeling of the training dataset using support of medical students team

#4
High F1-score value for the final pipeline – up to 80%

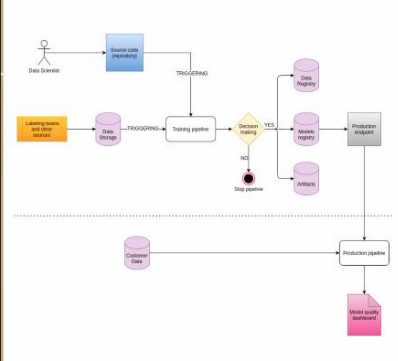
Unmute | Start Video | Participants (92) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

17:16 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Валерія, Kharkiv ITClu... | Павло Радюк | Daria Tamanova | Шпак Катерина

MLOPS SOLUTION



The diagram illustrates an MLOPS pipeline. It starts with 'Data ingestion' leading to 'Feature data preparation'. This feeds into 'Learning from data (training)'. A 'Data broker' also feeds into 'Training pipeline'. The 'Training pipeline' leads to a 'Decision engine'. From the 'Decision engine', there are two paths: 'Model registry' and 'Model deployment'. 'Model registry' leads to 'Production pipeline', which then feeds into 'Model monitoring dashboard'. 'Production pipeline' also feeds into 'Customer data', which then feeds into 'Model monitoring dashboard'. There are also 'Data lake' and 'Data warehouse' components connected to the 'Production pipeline'.

- Solution should support automated training, deploying mechanisms as well as saving model with artifacts into model registry.
- System should be activated by the model's code update or creation of a new version of the dataset.
- Prediction statistics are accumulating for the model monitoring mechanism.
- Developed to support multi cloud – AWS, GCP, Azure.

Unmute | Start Video | Participants (95) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

17:19 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | Валерія, Kharkiv ITClu... | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

WHY DATA SCIENTISTS DO MISTAKES?

<p>01 Lack of experience</p> <ul style="list-style-type: none"> The majority of mistakes exist due to the lack of experience with required technology. It can be not even juniors, because technology updates fast, and you usually should adapt to it. 	<p>02 Human factor</p> <ul style="list-style-type: none"> Fatigue, lack of motivation, personal problems are often the sources of a large number of mistakes.
<p>03 Laziness</p> <ul style="list-style-type: none"> Typical case that present in IT companies. Sometimes engineers prefer to do tasks in the last moment, that usually cause a lot of mistakes. 	<p>04 Weak management</p> <ul style="list-style-type: none"> A case that can exacerbate or solve most problems. Good management is important for any engineer.

Unmute | Start Video | Participants: 95 | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:22 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | Валерія, Kharkiv ITClu... | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

MISTAKES IN DATA SCIENCE

<p><input checked="" type="checkbox"/> Student's fever</p> <p>Category typical for Data Science freshmen.</p>	<p><input checked="" type="checkbox"/> Self-management issues</p> <p>Mistakes related to weak level of soft-skills.</p>
<p><input checked="" type="checkbox"/> Analysis and data preparation issues</p> <p>Mistakes related to weakness in initial data analysis.</p>	<p><input checked="" type="checkbox"/> Data Science solution issues</p> <p>DS project implementation level issues.</p>
<p><input checked="" type="checkbox"/> Coding issues</p> <p>Mistakes linked with code failures.</p>	

Unmute | Start Video | Participants: 94 | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:24 15.02.2023


Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | M | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

STUDENT'S FEVER

- STUDENT'S FEVER
- SOTA lover
Application the State of the art models for each task.
- Development hater
The opinion that the work of a data scientist is development Jupyter notebooks with experiments.
- Model designer
The opinion that to work as a Data Scientist you should know only modeling.
- Wheel inventor
The opinion that each solution should be implemented from scratch.



Unmute | Start Video | Participants: 95 | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:26 | 15.02.2023


Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | M | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

STUDENT'S FEVER

- STUDENT'S FEVER
- Young senior
The opinion that you can hold a Senior Data Science position without any experience.
- Gladiator
The opinion that competitions will make you a great commercial specialist.



Unmute | Start Video | Participants: 96 | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:28 | 15.02.2023


Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | M | Валерія, Kharkiv ITClu... | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

SELF-MANAGEMENT ISSUES

- Estimate shooter
Problem with estimating projects and separate tasks
- Business doesn't matter
Not taking into account business needs – only model quality
- Technical robot
Ignoring of the development of soft skills
- Distorting mirror
Implementation of the not required features due to the not understanding of the task definition



Unmute | Start Video | Participants (97) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:30 15.02.2023


Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | M | Валерія, Kharkiv ITClu... | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

ANALYSIS AND DATA PREPARATION ISSUES

- Blind analyst
Exploratory data analysis is performing without any visualization
- Broken faucet
Data leakage in the feature engineering stage
- Lost EDA
Building ML solution without any data analysis
- Garbage in – garbage out
Pushing data into a model without any preprocessing



Unmute | Start Video | Participants (97) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 2°C | 17:33 15.02.2023


Zoom Meeting You are viewing Kostiantyn Isaienkov's screen View Options

Tetiana Skrypnyk Kostiantyn Isaienkov Daria Tamanova M Валерія, Kharkiv ITClu... Павло Радюк Шпак Катерина

LIVE Custom Live Streaming Service

DATA SCIENCE SOLUTION ISSUES

- ML Golden Hammer
Attempts to apply the same model for each task.
- Black box user
Using model.fit() method without any setup and understanding.
- High score or nothing
Ignoring all Data Science lifecycle principles, just focusing on the accuracy improving.



Unmute Start Video Participants 94 Chat Share Screen Record Reactions Apps Whiteboards Leave

Пошук 2°C 17:35 15.02.2023


Zoom Meeting You are viewing Kostiantyn Isaienkov's screen View Options

Tetiana Skrypnyk Kostiantyn Isaienkov Daria Tamanova M Валерія, Kharkiv ITClu... Павло Радюк Шпак Катерина

LIVE Custom Live Streaming Service

CODING ISSUES

- Unit tests avoider
Ignoring unit testing in Data Science projects.
- Rock manuscripts
Code without any documentation.
- Hardcoder
Hard-coding constants, absolute paths, etc.
- Heap of everything
Ignoring project structure, all artifacts inside root directory.
- Fatal commit
Pushing into the repo all datasets and model artifacts.



Unmute Start Video Participants 93 Chat Share Screen Record Reactions Apps Whiteboards Leave

Пошук 0°C 17:38 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | M | Валерія, Kharkiv ITClu... | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

HOW TO START CAREER IN DATA SCIENCE

- Basic Knowledge**
 Good knowledge in Math, ML, and programming are mandatory.
- Understanding direction**
 Define and follow your goals.
- High motivation**
 Focus on growing yourself.
- Good CV**
 Check best practices to construct an attractive resume.
- Complete the interview**
 Get experience in interview completion

Unmute | Start Video | Participants (94) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 0°C | 17:41 | 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

Tetiana Skrypnyk | Kostiantyn Isaienkov | Daria Tamanova | M | Валерія, Kharkiv ITClu... | Павло Радюк | Шпак Катерина

LIVE Custom Live Streaming Service

DATA SCIENCE INTERNSHIP

1. **Paid 3-month duration program**
Usually internship program – 2 interns and mentor

2. **Real project during the internship**
Interesting R&D projects from scratch

3. **Plan for hard skills development**
Large Knowledge Evaluation system for DS

4. **Support of the team members**
Friendly team ready to help with working issues and share experience

5. **A job offer after the program completion**
More than 50% of the DS team – our interns in the past


Unmute | Start Video | Participants (91) | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Пошук | 0°C | 17:46 | 15.02.2023

Zoom Meeting | You are viewing Kostiantyn Isaienkov's screen | View Options

LET'S GET TO KNOW EACH OTHER

✉ k.isaienkov@quantummobile.com
 🌐 https://quantummobile.com/
 🏢 Quantum_Inc

More info about us 

Participants: 90 | Chat | Share Screen | Record | Reactions | Apps | Whiteboards | Leave

Windows Taskbar: Пошук, 0°C, 17:51, 15.02.2023

Zoom Meeting | Meeting Chat

Quantum LinkedIn <https://www.linkedin.com/company/10562076>
 Quantum Instagram https://www.instagram.com/quantum_inc/

Tigralwovna to Everyone 18:12
 Чи стикалися ви з задачами в рамках алгоритмів, досліджень пов'язаних з квантовим комп'ютигом?

Dmytro to Everyone 18:13
 200 баксів для Azure вистачить?

Роздають студентам

Дякую

Who can see your messages?

To: Everyone

Type message here...

Participants: 64 | Chat | Share Screen | Reactions | Apps | Whiteboards | More | Leave

Windows Taskbar: Пошук, -1°C, 18:16, 15.02.2023