⋮

Week 1 Discussion   13   62

Dave DeBarr

All Sections

Please post here any questions you might have during week 1.

| Search entries or author | Unread | ⬆ ⬇ | | ✓ **Subscribe** |

↩ **Reply**

○   **Dave DeBarr**   ⋮
    Apr 6, 2021

Okay, so I didn't read Rudy's latest reply. He needed the list on UW NetId identifiers for class members, in order to send out invitations to the Azure lab machines. I sent the list of UW NetId identifiers to Rudy a few minutes ago. He should have them set up by tomorrow. I will post an update tomorrow afternoon.

Edited by **Dave DeBarr** on Apr 6 at 10:08pm

↩ **Reply**   👍 (1 like)

○   **Dave DeBarr**   ⋮
    Apr 7, 2021

Everyone should have received their "Register for Lab – Deep Learning DSVM" invitation by now. You should be able to find it by visiting **https://gmail.uw.edu**

↩ **Reply**   👍

○   **Rajesh Arora**   ⋮
    Apr 7, 2021

Hi Dave, I have not received any email in my UW gmail account. I checked the spam folder as well.

↩ **Reply**   👍

○   **Chenwei Lin**   ⋮
    Apr 10, 2021

Hi Rajesh,

Do you still fail to receive e-mails from UW account? if so, please contact UW-IT (24x7)

https://itconnect.uw.edu/help/

Thanks,

Chenwei

↩ **Reply**   👍

○   **Rajesh Arora**   ⋮
    Apr 12, 2021

Thanks Chenwei. UW IT was able to help me, I am all set now.

↩ **Reply** 👍

Aeden Jameson
Apr 7, 2021

⋮

I did not receive the email.

↩ **Reply** 👍

Sandip Nath
Apr 7, 2021

⋮

I got it in my mailbox on **https://outlook.office365.com/mail/**　　signed in with sandipn@uw.edu

↩ **Reply** 👍

Aeden Jameson
Apr 8, 2021

⋮

I tried Sandip's suggestion using the email **aeden@uw.edu** and did NOT work. I received the error

Request Id: 0324b313-2e85-46b2-bccd-1c32d71c8e01
Correlation Id: 11b712c5-a00b-43c7-bad1-41779f070e9e
Timestamp: 2021-04-08T20:27:28Z
Message: AADSTS90020: The SAML 1.1 Assertion is missing ImmutableID of the user.

Perhaps that would be useful to the IT person that you're working with.

↩ **Reply** 👍

Dave DeBarr
Apr 8, 2021

⋮

As Sandip mentions above, it's possible you're using Office 365 instead of Gmail for accessing mail.

Rudy tells me you can find out whether you're using Office 365 or Gmail by visiting:

https://uwnetid.washington.edu/manage/?forward

↩ **Reply** 👍

Aeden Jameson
Apr 9, 2021

⋮

I tried signing in but get this error. Who do I need to call? This has gone on for several days now and I would like to get this resolved. The back and forth on the discussion board isn't working.

■■ Microsoft

Sorry, but we're having trouble signing you in.
AADSTS90020: The SAML 1.1 Assertion is missing ImmutableID of the user.
Troubleshooting details
If you contact your administrator, send this info to them.
Copy info to clipboard
Request Id: 6340512f-caa4-404f-86a8-142d99580e00
Correlation Id: 1f81d06d-4f15-48dd-a0f3-5117ee1db10c
Timestamp: 2021-04-10T00:24:47Z
Message: AADSTS90020: The SAML 1.1 Assertion is missing ImmutableID of the user.
Flag sign-in errors for review: Enable flagging

If you plan on getting help for this problem, enable flagging and try to reproduce the error within 20 minutes. Flagged events make diagnostics available and are raised to admin attention.

Edited by **Aeden Jameson** on Apr 9 at 5:29pm

↰ **Reply** 👍

Dave DeBarr
Apr 9, 2021   ⋮

Hi Aeden,

When are you encountering this error?

Let's suppose it's happening when you're trying to access Office 365. Is this because you're trying to access your email?

Did you navigate to **https://uwnetid.washington.edu/manage (Links to an external site.)**   to verify that your email is going to Office 365?

Thanks,

Dave

↰ **Reply** 👍

Aeden Jameson
Apr 9, 2021   ⋮

re: **https://uwnetid.washington.edu/manage/?forward** and "When are you encountering this error?"

I did check this and the emails were forwarded to **Aeden.jameson@gmail**. The email from rudy was in spam. I then did the following.

1. Click on "Register for the lab" in the email
2. I'm redirected to a login page where I enter "**aeden@uw.edu**" and my password
3. I then get the message posted above.

Does the email have to go to **aeden@uw.edu** and or do I need an Office 365 account?

These are the current settings.

Recovery:	**None established**
UW Office 365 ID:	**Sign up now**
UW G Suite ID:	**aeden@uw.edu**
UW Email:	Forwarded to **aeden.jameson@gmail.com**
MyUW.net Email:	Forwarded to **aeden.jameson@gmail.com**

[1]For security reasons, the office of UW Internal Audits recommends you change passwords every 90 days.

Screen Shot 2021-04-02 at 3.09.50 PM.png

↰ **Reply** 👍

SHOW MORE   2   2

Abhijit Majumdar
Apr 7, 2021   ⋮

Hello Everyone, is anyone having trouble downloading the kaggle.json file (with the API key)? When I click on the "Create New API Token" (within the Account tab), the json file is not downloading (although I am receiving the message that asks me to move the file into the Kaggle folder).

I believe the file is supposed to download into the Downloads folder, but I tried it on my Mac and my windows machine (using chrome, safari, brave and edge) without any luck.

Is anyone else facing this issue?

↰ **Reply**   👍 (1 like)

Dave DeBarr
Apr 7, 2021

If you clicked on the "Create New API Token" button using Edge on Windows, what happens if you click Control+J [press and hold the control key; press and release the 'j' key; then release the control key] and enter kaggle.json into the search box?

↰ **Reply**   👍

Sandip Nath
Apr 7, 2021

Same. There is nothing downloading. I changed browser, machine - no avail.

Ctrl-J download list does not have kaggle.json
Edited by **Sandip Nath** on Apr 7 at 6:36pm

↰ **Reply**   👍

Soo Lee
Apr 7, 2021

Same. I have trouble downloading, too.

↰ **Reply**   👍

Dave DeBarr
Apr 7, 2021

Sounds like your browser has a security setting that is preventing the download.

How would you feel about using fiddler to capture your key when it queries it from **www.kaggle.com/api/token?**

Maybe we have to setup a time where I rent a Windows VM for a few hours so folks can download their token?

I'll check with our sysadmin to see if he knows of a solution.

↰ **Reply**   👍

Dave DeBarr
Apr 7, 2021

It's possible that clicking the lock icon to the left of the url and selecting allow permissions will help, but I have my doubts if this is managed by an organization.

↰ **Reply**   👍

Dave DeBarr
Apr 8, 2021

It also seems possible that Kaggle may have just released a bug that affects the ability to create the kaggle.json file in your Downloads folder. If so, hopefully they will fix this soon, as it would be better to download larger datasets directly to your VM.

For now, perhaps the simplest alternative would be to download the data to your local host, then upload it to your VM.

After you have accepted the Kaggle invitation (using the URL in the assignment):

- navigate to **https://www.kaggle.com/c/ml530-2021-sp-mnist/data**

- scroll down to the bottom of the page
- click the "Download All" button

From a command line, use putty to upload the zip file to your VM:

"C:\Program Files (x86)\PuTTY\pscp.exe" -P YourPortNumber ml530-2021-sp-mnist.zip deeplearning@YourHostName

[pscp (PuTTY's secure file copy) could be located in "C:\Program Files\PuTTY\pscp.exe"]

After you have created predictions.csv on your VM, use putty to download the csv file from your VM:

"C:\Program Files (x86)\PuTTY\pscp.exe" -P YourPortNumber deeplearning@YourHostName:/path/to/predictions.csv predictions.csv

Navigate back to **https://www.kaggle.com/c/ml530-2021-sp-mnist**   and use the "Submit Predictions" button to upload your predictions.

↰ **Reply** 👍

Benjamin Johnson
Apr 8, 2021

I just tried, and the download seems to be working now (it was not working earlier this morning).

↰ **Reply** 👍 (1 like)

Abhijit Majumdar
Apr 9, 2021

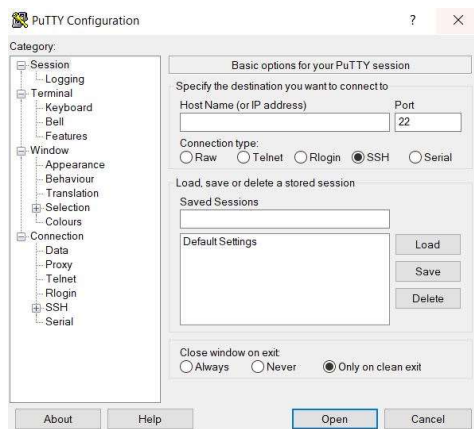Good Afternoon Dave, Benjamin is correct, the download is working now :-) ! I was able to fetch my API token :-).

↰ **Reply** 👍

Soo Lee
Apr 7, 2021

Hi

Where can I find the Host name for Putty?

↰ **Reply** 👍

Dave DeBarr
Apr 7, 2021

Click the funky looking connector on the card for the VM and choose "Connect via ssh": the port number is the -p parameter

↰ **Reply** 👍

Soo Lee
Apr 7, 2021

I figured out what the port number is. Thanks!!!

↩ **Reply**   👍

Sandip Nath
Apr 7, 2021

I created a new kaggle account using the uw account. I had joined the competition and tried to get the kaggle.json. Everything was fine till for some reason I decided to verify my phone number.

Kaggle detected that it is a duplicate phone number, matching with that from my existing Kaggle account (based on a gmail address), and instantly blocked the uw account.

I have put a request for support but have serious doubts that it will be resolved any time soon.

Edited by **Sandip Nath** on Apr 7 at 9:52pm

↩ **Reply**   👍

Dave DeBarr
Apr 7, 2021

If you can't use your UW account, just use your other account.

↩ **Reply**   👍 (1 like)

Rajesh Arora
Apr 8, 2021

Hi Chenwei, Dave, I have an issue with my UW email account (rarora2@uw.edu). For some reason no emails are coming to this account. When you send something to this account, you get a delivery failure. However, I can logon into this account and send emails. For this reason I did not get the invitation to access Azure lab machine. Is there way for me to get in touch with UW tech folks to get this resolved?

Edited by **Rajesh Arora** on Apr 8 at 6:33pm

↩ **Reply**   👍

Rajesh Arora
Apr 9, 2021

Hi Chenwei, Dave, I called UW tech help to get my email account fixed. Since it is working now, would it be possible for you to resend the invitation to get Azure access?

↩ **Reply**   👍

Dave DeBarr
Apr 9, 2021

I asked Rudy to resend the invitation.

Dave

↩ **Reply**   👍

Rajesh Arora
Apr 12, 2021

Thanks Dave. I have received the invitation, I am all set now. Thanks again for your help.

↩ **Reply**   👍

Eric Kuo

Apr 8, 2021

Hi Dave, Do you have step-by-step instructions for setting everything up? If so, can you post them?

↩ **Reply** 👍

Dave DeBarr

Apr 8, 2021

Hi Eric, I was hoping the notes in the assignment (_____**Assignment 1**_____) would be sufficient? Is something unclear? Thanks, Dave

↩ **Reply** 👍

Eric Kuo

Apr 8, 2021

Thanks Dave! I didn't think to check the assignment.

↩ **Reply** 👍

Carl Gaither

Apr 8, 2021

I think there's a missing "-c" in the Assignment 1 instructions for downloading the mnist data into the VM. I kept getting a 401-Unauthorized error until I added the -c.

Specifically, the assignment instructions say

kaggle competitions download ml530-2021-sp-mnist

but you should do

kaggle competitions download -c ml530-2021-sp-mnist.

At least this worked for me. I'm running the code now. We'll see how things work.

↩ **Reply** 👍

Dave DeBarr

Apr 8, 2021

Funky. It worked as it is for me. I'm using Kaggle API version 1.5.12 with Python version 3.8.5. Which version are you using?

$ kaggle competitions download ml530-2021-sp-mnist

Downloading ml530-2021-sp-mnist.zip to /home/dadebarr

  99%|████████████████████████████████| 31.0M/31.3M [00:02<00:00, 11.7MB/s]

100%|████████████████████████████████| 31.3M/31.3M [00:02<00:00, 11.5MB/s]

$ kaggle --version

Kaggle API 1.5.12

$

↩ **Reply** 👍

Carl Gaither

Apr 9, 2021

I'm using Kaggle API 1.5.12 (checked with kaggle --version) and whatever Python is installed in the VM. But at least everything worked. At some point during the course it will become clear what's going on.

↩ **Reply** 👍

Aeden Jameson

Apr 9, 2021

I have a general question, why is the content not organized, imho, as clearly as the other classes have been? For example, below is an example of the last class and every other class I've taken. The content and topics were clearly laid out and there was no need to visit another site. Just my two cents.

MLEARN 520 A Wi 21: Advanced Machine Learning ⌄

📊 View Course Stream

📅 View Course Calendar

🔔 View Course Notifications

To Do

Nothing for now

Recent Feedback

✓ End-of-Course Survey
1 out of 1

✓ Lesson 09 Knowledge Check
3 out of 5

✓ Lesson 09 Discussion
5 out of 5

2 more ...

Welcome & Get to Know Your Course

→ Course Overview	📑 Course Resources	👥 Course Team	👥 Course Objectives

Part 1	Part 2	Part 3	Part 4
Week 1 Introduction to Advanced Machine Learning	**Week 3** Random Forest & Gradient Boosted Trees	**Week 6** Unsupervised Learning	**Week 9** Forecasting
Week 2 Decision Tree	**Week 4** Support Vector Machine (SVM)	**Week 7** Natural Language Processing	**Week 10** Building Machine Learning Applications
	Week 5 Stacking and Blending	**Week 8** Recommendation	

Lesson 8: Mining Unstructured Data (Text Mining and Natural Language Processing)

↩ **Reply** 👍 (2 likes)

Dave DeBarr

Apr 9, 2021

Hi Aeden,

For what it's worth: I'm sorry you don't feel the course content is well organized.

I did cover an analogous outline when reviewing the syllabus with the class.

Please contat **help@uw.edu** to work out the login issue.

Thanks,

Dave

↩ **Reply** 👍

Luis Garduno
Apr 10, 2021

While I agree with you that having the topic for each week laid out within the course's Canvas home page would be nice, I think the overview of topics for each week were discussed during the syllabus overview:

1. Review of machine learning basics and the extension to multi-layer perceptrons ("deep" means many layers)
2. Introduction: Chapter 1 and 2 from DLP
3. Fundamentals: Chapters 3 and 4 from DLP
4. ConvNets: Chapters 1 and 10 from DLI
5. ConvNets: Chapter 5 from DLP [transfer learning part one]
6. Embeddings and Sequences: Chapters 2 and 11 from DLI
7. Embeddings and Sequences: Chapter 6 from DLP
8. Transformers and Sequence-to-Sequence Models [transfer learning part two]
9. Generative Models: Chapters 3 and 12 from DLI; Chapter 8 from DLP
10. Deep Reinforcement Learning: Chapters 4 and 13 from DLI

And comparing it to the previous course, I honestly value much more that the instructor is engaged in the discussions after class.

↩ **Reply** 👍

Aeden Jameson
Apr 11, 2021

My observation wasn't merely about whether some table of contents was covered and put somewhere at some point. Every other class, that I've taken at least, has managed to have the weeks also reasonably and consistently organized within canvas. I don't need to personally like it, but consistency of the logistics and content organization goes a long way week over week and class over class when time is a premium; work, class and whatever other responsibilities one may have. All the logistics take away from the value. In addition to the initial deviation in content organization the link to the zoom meeting was in some pdf and then I received a meeting id that didn't work. Every other class has managed to have the link in the Zoom section of the course.

None the less, I'm sure it'll be fine and we'll all get through it.

Edited by **Aeden Jameson** on Apr 11 at 10:22am

↩ **Reply** 👍 (1 like)

Luis Garduno
Apr 11, 2021

Yep, I agree that courses should try to have the same or similar format/logistics. What I've seen in general with all of the courses in this certificate is that it doesn't seem like they had the input from an instructional designer (or maybe they had initially and were not reviewed again), consistency in their format would just be one thing, but there are other areas that I think are missing. Anyways, I've passed some of my feedback on the previous courses to the certificate's program manager, Piyu Roy (piyuroy@uw.edu), in case you want to share yours.

↩ **Reply** 👍

Siri Sadashiva
Apr 10, 2021

Hi Dave/Chenwei,

I still have not received any email (opened the outlook and gmail). I have set the UW **Forward to UW Office 365.**

Please let me know whom to contact

Thank you

Edited by **Siri Sadashiva** on Apr 10 at 5:28pm

↩ **Reply** 👍

Aeden Jameson
Apr 11, 2021

I did get access finally so from my experience so far it doesn't matter where you receive the email. So wherever you were receiving email you may want to check the spam folder unless of course you previous had forwarding set to "stopped." You will at least want to be sure that UW Office 365 ID is enabled under **https://uwnetid.washington.edu/manage/?settings** . This appears to be required in order for the SSO process between Microsoft and the UW to work.

Edited by **Aeden Jameson** on Apr 11 at 10:36am

↩ **Reply** 👍

Glen A Frison
Apr 11, 2021

How do we turn in the answer to the question in the homework?

↩ **Reply** 👍

Dave DeBarr
Apr 11, 2021

In the upper, righthand corner of the assignment, you should see a button labeled "Start Assignment." You should then see tabs at the bottom of the assignment labeled "File Upload" and "Text Entry". Either should be fine.

↩ **Reply** 👍

Andrew Feldman
Apr 12, 2021

Sorry, quick clarification - there are two components, the several-part Question 1 & the Kaggle task in which we are asked to ultimately upload our predictions to Kaggle. I have (1) uploaded a PDF of my solution to the several-part Question 1 using the "File Upload" button as you described, and (2) submitted my predictions.csv to Kaggle for the second part. Is that it? Do I need to also submit my predictions.csv using the "File Upload" button? Am I done once I see myself on the Kaggle leaderboard?

↩ **Reply** 👍

Dave DeBarr
Apr 12, 2021

There is no need to upload predictions.csv to canvas. Please just include your display name in what you upload to canvas. Thanks!

↩ **Reply** 👍

Luis Garduno
Apr 11, 2021

I did not have the registering issues that other people had, but I forgot to stop the VM after first setting everything up and now I cannot use it :\

I'm waiting to get help from Rudy, but I just wanted to share it so you don't make the same mistake, make sure you stop your VM when not using it :)

↩ **Reply** 👍

Andrew Feldman
Apr 12, 2021

Made this mistake too! To cover my needs in the meantime, I set up an environment that could run the lab in docker on an Ubuntu 18.04 VM

docker run -it afeld2012/anfeldml:kaggleml_share

I have a VM with GPU support and CUDA 10 on the host VM, so I actually run it with

docker run -it --gpus all afeld2012/anfeldml:kaggleml_share

↩ **Reply** 👍 (1 like)

Andrew Feldman
Apr 12, 2021

Also FYI, if you try to set up your own GPU enabled VM on Azure or AWS - there is a long approval time because people abuse GPU VMs for cryptocurrency mining and (presumably) the number of such SKUs is limited :) On both providers, the limitation is expressed as "your account has a quota of 0 cores for this GPU-enabled SKU". You need to file a support request to get the quota for that particular SKU increased to more than 0.

On Azure, the support ticket was resolved within an hour and I was able to get an NC SKU (NVIDIA GPU-enabled SKU) and run the above docker image. On AWS I still haven't received approval in spite of a stern conversation with customer support ;) That's just my experience.

↩ **Reply** 👍

Andrew Feldman
Apr 12, 2021

Of course Rudy may provide a faster and lower-effort resolution for you haha :)
Edited by **Andrew Feldman** on Apr 12 at 4:33am

↩ **Reply** 👍

Luis Garduno
Apr 12, 2021

Thanks for the help tips, Andrew! Rudy was able to add a few hours to my VM.

↩ **Reply** 👍

Carl Gaither
Apr 12, 2021

I'm glad I'm not the only one who forgot - after I ate up my entire 50 hours.

↩ **Reply** 👍 (1 like)

Andrew Feldman
Apr 12, 2021

So, running this code I see that many train/test runs are happening one-after-another, and each run may have a different number of epochs. I guess the change in epoch count is a result of keras-tuner optimizing the exploration/exploitation tradeoff?

↩ **Reply** 👍

Dave DeBarr
Apr 12, 2021

We'll review in class, but as shown in slide 41 from last week's class: there are 62 total hyperparameter configurations evaluated across 90 trials, with early stopping for candidates that do not perform well. 62 is the sum of the number of candidates in round 0 of the four brackets.

↩ **Reply** 👍 (1 like)

○   **Hans-Erik Andersen**
            Apr 13, 2021

Is there a graded discussion for this week? I'm not able to find it.

← **Reply**   👍 (1 like)

○   **Dave DeBarr**
            Apr 13, 2021

No graded discussion.

← **Reply**   👍 (2 likes)

○   **Alex Alimov**
            Apr 13, 2021

Hey Dave,

I know I asked a version of this question in-class, but I was curious if you can give any additional feedback. What types of ML/DS positions is it reasonable to apply for after completing this program? Has anything that may be crucial during a job search been left out of the program? If so, what else should be thoroughly covered?

Thank you,

Alex

Edited by **Alex Alimov** on Apr 13 at 4:56pm

← **Reply**   👍 (1 like)

○   **Dave DeBarr**
            Apr 13, 2021

I *think* you should be able to get any job where the goal is to "solve a kaggle-like task"; i.e. given a data set and a model evaluation metric, I would expect you to be able to solve the problem with an approximation to a state-of-the-art solution in a "reasonable" amount of time [where "reasonable" depends on your familiarity with the resources used in our work environment].

Look at the currently active kaggle competitions: **Kaggle Competitions**

Those are real tasks, where folks are willing to pay some money to get a solution.

If folks needed new research to be able to solve their task, chances are decent that they won't be posting their new research problem to kaggle. If you'd like to get a job developing new learning methods (i.e. research), you're probably going to need the PhD degree; but if you just want to apply state-of-the-art (existing solutions that make it to the "top of the leaderboard"), then having a master's degree or a certificate is probably sufficient [with the master's degree candidates perhaps having a bit of an edge over the certificate candidates, with the trade-off being the master's degree candidates also probably cost a bit more].

When I interview someone, I typically want to understand why the change in direction [if you've been working in some other problem space]; and I'll ask detailed questions to gauge your interest in understanding/applying state of the art solutions.

Usually I only get to ask 3 questions:

1) walk me through a project where you created and evaluated a machine learning model: I'm going to ask lots of questions about why you made the decisions that you made, and how well you explored alternatives.

2) a simple coding question, usually something involving estimating uncertainty for the evaluation metric from the previous problem: I just don't want to be surprised when you finally join the team and say "Coding? I don't do coding." I am willing to provide guidance on how to do it. My expectations for your knowledge will depend on the "level" of the position you're interviewing for.

3) literally I'll pose a kaggle question [probably]: I want to know what it would be like to brainstorm a solution to a realistic-ish problem with you.

Note: some folks ask funky coding questions during the interview process (e.g. **LeetCode - The World's Leading Online Programming Learning Platform**   "medium" problems). For what it's worth: I've never had to solve the longest common subsequence problem at work, but

the mix of computer science folks involved in our field means some folks may want to test your coding/algorithm skills too.

Edited by **Dave DeBarr** on Apr 13 at 11:46pm

↩ **Reply**   👍 (2 likes)

Alex Alimov

May 21, 2021

Thank you very much!

↩ **Reply**   👍