

Generalise for unseen positions in date

In any instruments (cheques, forms etc) the date is a mandatory field, and the same needs to be extracted with maximum accuracy to automate use cases related to it.

In the real-world scenario, we receive data which is biased positionally, for example the **month** would be only from 01-12, hence we don't have training data that covers all the possible combinations across all positions of DDMMYYYY.

But while inference the model should not be biased towards certain digits across positions as it may bypass many invalid dates

The training and test dataset is specially synthesized to test the generalisation of model across positions, as some of the positions in training data is heavily biased for certain digit/digits

To submit the results, one would require the credentials that are used to log in to the RDP instance.

```
sub_json = sub_df.to_json()
data = {'username': 'deepak', 'password': 'xxxx', 'submission': sub_json, 'tag': 'test2'}
r = requests.post(SUBMISSION_URL, json=data)
```

```
r.text
```

```
'{"success":true,"score":0.0}'
```

Link to download required data - <http://13.234.225.243:9600>

← → ↻ 🏠 ⚠ Not secure | 13.234.225.243:9600

Directory listing for /

- [sample_submission.csv](#)
- [test_data.tar](#)
- [train_data.csv](#)
- [train_data.tar](#)

The training and test images can be extracted from the respective tar files.