PROJECT 2, GALAXY REDSHIFT ESTIMATION

Farah Najla - IVC18 IVC on Astrostatistics and Machine Learning Day 9 Hands-on

Import data

In [2]:
```python
import numpy as np
import matplotlib.pyplot as plt
import astropy.io.fits as fits

data_url = 'https://anirut.space/data/decals_galaxy.fits'
hdul = fits.open(data_url)
data = hdul[1].data
```

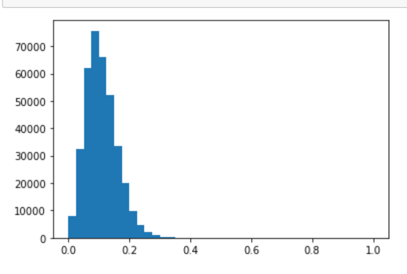
In [3]:
```python
# Check header
hdul[1].header
```

Out[3]:
```
XTENSION= 'BINTABLE'           / binary table extension
BITPIX  =                    8 / 8-bit bytes
NAXIS   =                    2 / 2-dimensional table
NAXIS1  =                   32 / width of table in bytes
NAXIS2  =               368407 / number of rows in table
PCOUNT  =                    0 / size of special data area
GCOUNT  =                    1 / one data group
TFIELDS =                    6 / number of columns
EXTNAME = 'Joined  '           / table name
TTYPE1  = 'RA      '           / label for column 1
TFORM1  = 'D       '           / format for column 1
TTYPE2  = 'DEC     '           / label for column 2
TFORM2  = 'D       '           / format for column 2
TTYPE3  = 'MAG_G   '           / label for column 3
TFORM3  = 'E       '           / format for column 3
TTYPE4  = 'MAG_R   '           / label for column 4
TFORM4  = 'E       '           / format for column 4
TTYPE5  = 'MAG_Z   '           / label for column 5
TFORM5  = 'E       '           / format for column 5
TTYPE6  = 'spec_z  '           / label for column 6
TFORM6  = 'E       '           / format for column 6
DATE-HDU= '2021-08-03T11:01:35' / Date of HDU creation (UTC)
STILVERS= '3.4-1   '           / Version of STIL software
STILCLAS= 'uk.ac.starlink.votable.FitsPlusTableWriter' / STIL Author class
```

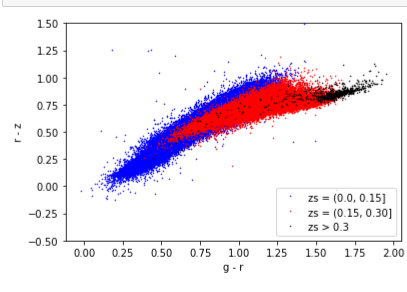
In [4]:
```python
# Parameters
g = data["MAG_G"]
r = data["MAG_R"]
z = data["MAG_Z"]
zs = data["spec_z"]
g_r = g - r
r_z = r - z
g_z = g - z
```

Plot the relations: g,r,z and redshift

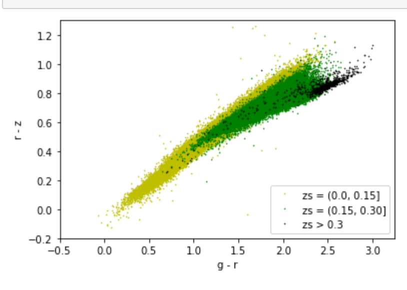
In [5]:
```python
plt.hist(zs, bins=40)
plt.show()
```



In [6]:
```python
# Color-color diagram
# g-r vs r-z
plt.plot(g_r[(zs > 0.0)  & (zs <= 0.15)], r_z[(zs > 0.0)  & (zs <= 0.15)], 'b.',
         markersize=1, label="zs = (0.0, 0.15]")
plt.plot(g_r[(zs > 0.15)  & (zs <= 0.30)], r_z[(zs > 0.15)  & (zs <= 0.30)], 'r.',
         markersize=1, label="zs = (0.15, 0.30]")
plt.plot(g_r[zs > 0.3], r_z[zs > 0.3], 'k.',
         markersize=1, label="zs > 0.3")
plt.ylim(-0.5, 1.5)
plt.xlabel("g - r")
plt.ylabel("r - z")
plt.legend(loc=4)
plt.show()
```



In [13]:
```python
# Color-color diagram
# g-z vs r-z
plt.plot(g_z[(zs > 0.0)  & (zs <= 0.15)], r_z[(zs > 0.0)  & (zs <= 0.15)], 'y.',
         markersize=1, label="zs = (0.0, 0.15]")
plt.plot(g_z[(zs > 0.15)  & (zs <= 0.30)], r_z[(zs > 0.15)  & (zs <= 0.30)], 'g.',
         markersize=1, label="zs = (0.15, 0.30]")
plt.plot(g_z[zs > 0.3], r_z[zs > 0.3], 'k.',
         markersize=1, label="zs > 0.3")
plt.ylim(-0.2, 1.3)
plt.xlim(-0.5, 3.25)
plt.xlabel("g - r")
plt.ylabel("r - z")
plt.legend(loc=4)
plt.show()
```



Setting up neural network layers

In [14]:
```python
# Input magnitude 6 columns
object_magnitude = np.stack((g, r, z, g_r, r_z, g_z), axis=-1)

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(object_magnitude, zs, test_size=0.2)
```

In [15]:
```python
import tensorflow as tf
from tensorflow.keras.layers import Dense, Dropout
model = tf.keras.Sequential()
```

In [16]:
```python
model.add(Dense(16, input_dim=6, activation='relu'))
model.add(Dense(16, activation='relu'))
model.add(Dense(1))
```

Selecting optimizer and loss function

In [17]:
```python
model.compile(optimizer=tf.keras.optimizers.Adam(0.001), loss="mse")
model.summary()
```

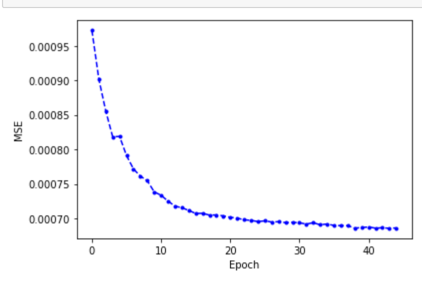
```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 16)                112
_________________________________________________________________
dense_1 (Dense)              (None, 16)                272
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 17
=================================================================
Total params: 401
Trainable params: 401
Non-trainable params: 0
_________________________________________________________________
```

Training the neural networks

In [18]:
```python
# 50 epochs
save_log = model.fit(x_train, y_train, epochs=50, batch_size=64)
```

```
Epoch 1/50
4606/4606 [==============================] - 14s 3ms/step - loss: 0.0552
Epoch 2/50
4606/4606 [==============================] - 10s 2ms/step - loss: 0.0011
Epoch 3/50
4606/4606 [==============================] - 7s 2ms/step - loss: 0.0011
Epoch 4/50
4606/4606 [==============================] - 7s 2ms/step - loss: 0.0011
Epoch 5/50
4606/4606 [==============================] - 7s 1ms/step - loss: 0.0010
Epoch 6/50
4606/4606 [==============================] - 7s 2ms/step - loss: 9.7249e-04A: 0s - loss
Epoch 7/50
4606/4606 [==============================] - 9s 2ms/step - loss: 9.0181e-04
Epoch 8/50
4606/4606 [==============================] - 11s 2ms/step - loss: 8.5570e-04
Epoch 9/50
4606/4606 [==============================] - 10s 2ms/step - loss: 8.1828e-04
Epoch 10/50
4606/4606 [==============================] - 9s 2ms/step - loss: 8.1871e-04A: 0s -
Epoch 11/50
4606/4606 [==============================] - 8s 2ms/step - loss: 7.9103e-04
Epoch 12/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.7175e-04A
Epoch 13/50
4606/4606 [==============================] - 10s 2ms/step - loss: 7.6117e-04: 0s - loss: 7.6155e-
Epoch 14/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.5465e-04
Epoch 15/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.3845e-04
Epoch 16/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.3349e-04
Epoch 17/50
4606/4606 [==============================] - 8s 2ms/step - loss: 7.2503e-04
Epoch 18/50
4606/4606 [==============================] - 11s 2ms/step - loss: 7.1776e-04: 0s - l
Epoch 19/50
4606/4606 [==============================] - 12s 3ms/step - loss: 7.1540e-04
Epoch 20/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.1159e-04
Epoch 21/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.0709e-04
Epoch 22/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.0733e-04
Epoch 23/50
4606/4606 [==============================] - 8s 2ms/step - loss: 7.0476e-04
Epoch 24/50
4606/4606 [==============================] - 10s 2ms/step - loss: 7.0504e-04
Epoch 25/50
4606/4606 [==============================] - 8s 2ms/step - loss: 7.0363e-04
Epoch 26/50
4606/4606 [==============================] - 9s 2ms/step - loss: 7.0155e-04
Epoch 27/50
4606/4606 [==============================] - 8s 2ms/step - loss: 7.0040e-04
Epoch 28/50
4606/4606 [==============================] - 8s 2ms/step - loss: 6.9794e-04
Epoch 29/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.9713e-04
Epoch 30/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9536e-04
Epoch 31/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.9643e-04
Epoch 32/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9475e-04
Epoch 33/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9522e-04
Epoch 34/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9381e-04
Epoch 35/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9422e-04
Epoch 36/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.9364e-04
Epoch 37/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.9173e-04
Epoch 38/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9373e-04
Epoch 39/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.9071e-04
Epoch 40/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.9187e-04
Epoch 41/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8977e-04
Epoch 42/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8946e-04
Epoch 43/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8934e-04
Epoch 44/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8596e-04
Epoch 45/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8669e-04
Epoch 46/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8756e-04
Epoch 47/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8584e-04
Epoch 48/50
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8643e-04
Epoch 49/50
4606/4606 [==============================] - 13s 3ms/step - loss: 6.8562e-04
Epoch 50/50
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8621e-04
```

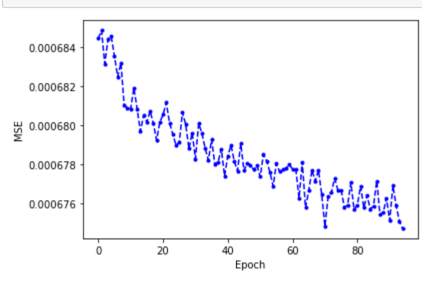
In [19]:
```python
plt.plot(save_log.history['loss'][5:], "b.--")
plt.xlabel("Epoch")
plt.ylabel("MSE")
plt.show()
```



In [21]:
```python
# 100 epochs
save_log = model.fit(x_train, y_train, epochs=100, batch_size=64)
```

```
Epoch 1/100
4606/4606 [==============================] - 13s 3ms/step - loss: 6.8410e-04
Epoch 2/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8493e-04
Epoch 3/100
4606/4606 [==============================] - 12s 3ms/step - loss: 6.8542e-04
Epoch 4/100
4606/4606 [==============================] - 12s 3ms/step - loss: 6.8355e-04
Epoch 5/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8457e-04A: 0s - loss: 6.
Epoch 6/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.8444e-04
Epoch 7/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.8486e-04
Epoch 8/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8314e-04
Epoch 9/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8441e-04
Epoch 10/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8455e-04
Epoch 11/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8354e-04
Epoch 12/100
4606/4606 [==============================] - 13s 3ms/step - loss: 6.8248e-04: 1s -
Epoch 13/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.8318e-04
Epoch 14/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.8100e-04
Epoch 15/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.8088e-04
Epoch 16/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.8084e-04
Epoch 17/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.8191e-04
Epoch 18/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.8084e-04
Epoch 19/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7968e-04
Epoch 20/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8049e-04
Epoch 21/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8015e-04: 0s -
Epoch 22/100
4606/4606 [==============================] - 12s 2ms/step - loss: 6.8072e-04: 1s - loss: 6.8300e-0 -
ETA: 1s - loss - E
Epoch 23/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8009e-04
Epoch 24/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.7923e-04: 1
Epoch 25/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8014e-04
Epoch 26/100
4606/4606 [==============================] - 14s 3ms/step - loss: 6.8054e-04
Epoch 27/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8115e-04
Epoch 28/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8011e-04
Epoch 29/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7955e-04
Epoch 30/100
4606/4606 [==============================] - 12s 3ms/step - loss: 6.7897e-04
Epoch 31/100
4606/4606 [==============================] - 13s 3ms/step - loss: 6.7910e-04
Epoch 32/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.8066e-04
Epoch 33/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.8002e-04
Epoch 34/100
4606/4606 [==============================] - 12s 3ms/step - loss: 6.7879e-04
Epoch 35/100
4606/4606 [==============================] - 15s 3ms/step - loss: 6.7961e-04
Epoch 36/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7827e-04
Epoch 37/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.8009e-04
Epoch 38/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7957e-04
Epoch 39/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7882e-04
Epoch 40/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7819e-04A: 4s - ETA: 0s - loss:
Epoch 41/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7926e-04
Epoch 42/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7798e-04A: 0s - loss: 6.77
Epoch 43/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7811e-04
Epoch 44/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7878e-04A - ETA: 0s -
Epoch 45/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7739e-04
Epoch 46/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7842e-04
Epoch 47/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7897e-04
Epoch 48/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.7813e-04: 0s - loss: 6.
Epoch 49/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7765e-04
Epoch 50/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7904e-04
Epoch 51/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7766e-04
Epoch 52/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7805e-04
Epoch 53/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7796e-04
Epoch 54/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7771e-04
Epoch 55/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7793e-04
Epoch 56/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7735e-04
Epoch 57/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7849e-04
Epoch 58/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7817e-04
Epoch 59/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7758e-04A: 0
Epoch 60/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7686e-04A: 0s - loss: 6
Epoch 61/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7803e-04
Epoch 62/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7762e-04
Epoch 63/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7773e-04
Epoch 64/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7779e-04A: 0s -
Epoch 65/100
4606/4606 [==============================] - 12s 3ms/step - loss: 6.7796e-04
Epoch 66/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7774e-04
Epoch 67/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7774e-04A:
Epoch 68/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7626e-04
Epoch 69/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7811e-04
Epoch 70/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7576e-04
Epoch 71/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7664e-04
Epoch 72/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7769e-04
Epoch 73/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7713e-04
Epoch 74/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.7769e-04
Epoch 75/100
4606/4606 [==============================] - 12s 3ms/step - loss: 6.7644e-04
Epoch 76/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7481e-04
Epoch 77/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7637e-04
Epoch 78/100
4606/4606 [==============================] - 11s 2ms/step - loss: 6.7656e-04
Epoch 79/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7727e-04
Epoch 80/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.7666e-04
Epoch 81/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7666e-04
Epoch 82/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7579e-04
Epoch 83/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7589e-04
Epoch 84/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7705e-04
Epoch 85/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7567e-04
Epoch 86/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7588e-04
Epoch 87/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7687e-04
Epoch 88/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7577e-04
Epoch 89/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7639e-04
Epoch 90/100
4606/4606 [==============================] - 10s 2ms/step - loss: 6.7570e-04
Epoch 91/100
4606/4606 [==============================] - 9s 2ms/step - loss: 6.7585e-04
Epoch 92/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7714e-04
Epoch 93/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7541e-04
Epoch 94/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7551e-04
Epoch 95/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7625e-04
Epoch 96/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7510e-04
Epoch 97/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7690e-04
Epoch 98/100
4606/4606 [==============================] - 8s 2ms/step - loss: 6.7586e-04A: 4s - loss:
Epoch 99/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7505e-04
Epoch 100/100
4606/4606 [==============================] - 7s 2ms/step - loss: 6.7470e-04
```

In [22]:
```python
plt.plot(save_log.history['loss'][5:], "b.--")
plt.xlabel("Epoch")
plt.ylabel("MSE")
plt.show()
```

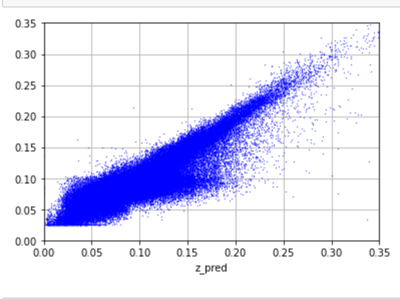


Predict the redshift and find the standard deviation

In [23]:
```python
y_pred = model.predict(x_test)[:, 0]
sigma = np.std((np.abs(y_pred - y_test)) / (1.0 + y_test))
print("S.D. = ", sigma)
```

```
S.D. =  0.015480323
```

In [24]:
```python
plt.xlim(0.0, 0.35)
plt.ylim(0.0, 0.35)
plt.plot(y_test, y_pred, 'b.', markersize = 0.5)
plt.xlabel("z_spec")
plt.xlabel("z_pred")
plt.grid()
plt.show()
```



In [25]:
```python
model.save("model.hdf5", overwrite=True, include_optimizer=True,
save_format=None, signatures=None, options=None)
```