Find the following entities from the given resumes using OCR–

1. Mobile number
2. Email id
3. Any hyperlink (if in documents)
4. College Qualification (e.g MCA, B.tech etc)

**Pdf 1 result:**

**Mobile :** ['8966786540', '9634857642']

**Email:** ['234@gmail.com']

**Hyperlink**: [ ]

**Education :** ['BLE','ME','cia', 'cia', 'BS', 'BS','cha', 'BS', 'cia', 'X''cia','X', 'cha', 'Bachelor of Engineering', 'cha', 'CLAR', 'cla','cha']

**Pdf 2 result:**

**Mobile :** ['9876543210']

**Email:** ['Gargpurvi2022@gmail.com']

**Hyperlink :** []

**Education :** ['X','cta','CIAT','BUS','ME','X','BUS','ME','X','cha','MAS']

**Pdf 3 result :**

**Mobile :** ['347-123-739']

**Email:** ['waleedahmad98@gmail.com']

**Hyperlink:** [ ]

**Education :** ['X', 'cia', 'cia', 'cia', 'cha', 'cra', 'cra', 'MS', 'cha', 'MLS', 'MyS']

**Pdf 4 result :**

**Mobile :** ['9628212345']

**Email:** ['amitkumar1234S@gmail.com']

**Hyperlink:** [ ]

**Education :** ['BJE', 'cha', 'X', 'mca', 'X', 'MyS', 'mca', 'MCA', 'BCA', 'MyS','MS', 'mca', 'MS', 'cia']

**NOTE:**

Mobile extraction works fine, email id also extracted except that first pdf case.

There is no hyperlink in pdf, however it would work

**There is chance in improvement in education degree extraction, I have extracted some degree along with wrong information. Still needs to improve.**