

PERCEPTRON *v 1.0.0.0*

A Next Generation Top-Down Proteoform
Identification and Characterization Platform

USER MANUAL

Biomedical Informatics Research Laboratory, Department of Biology
Lahore University of Management Sciences
<http://birl.lums.edu.pk/>

Last Modified: 20-12-2019

1. Table of Contents

1. Table of Contents	2
2. Table of Figures	4
3. Introduction to PERCEPTRON	6
3.1. About PERCEPTRON	6
3.2. Features	6
4. Hardware and Software.....	8
4.1. Hardware	8
4.2. Software	8
4.3. Testing.....	8
5. Video Tutorials	9
6. Getting Started with PERCEPTRON.....	9
7. GUI Description.....	11
7.1 Window 1: PERCEPTRON Tool for Top-down Proteomics.....	11
7.2 Window 2: Protein Search Query.....	12
7.2.1 Basic Parameters.....	12
7.2.2 Set Experimental Parameters	13
7.2.3 Set <i>De Novo</i> Sequencing Parameters	14
7.2.4 Set Protein Modifications Parameters.....	15
7.2.5 Set Scoring Components Weight	16
7.3 Window 3: Summary and Detailed Results View.....	17
8. Search.....	18
8.1 File Formats.....	18
8.1.1 Raw to mzXML File Format Conversion	18

8.1.2	Raw to mzML File Format Conversion	18
8.1.3	MzXML to MGF File Format Conversion	19
8.1.4	MGF to Flat Text File Format Conversion	19
8.2	Parameters	20
8.3	Databases.....	20
8.4	Modes	20
9.	Deploying PERCEPTRON	21
9.1	Installing Visual Studio Code (VS Code)	21
9.2	Installing SQL Server Management Studio (SSMS)	23
9.3	Installing Node.js.8.11.1	24
9.4	Installing SQL 2017	26
9.5	Installing Angular 1.7.4.....	27
9.6	Importing files into SQL	29
10.	Building PERCEPTRON	30
10.1	PERCEPTRON Application Programming Interface (API)	30
10.2	PERCEPTRON Front-End	34
10.3	Connecting PERCEPTRON API to PERCEPTRON Front-End.....	35
10.4	PERCEPTRON Local Service.....	38
11.	References.....	41

2. Table of Figures

Figure 1. PERCEPTRON Homepage and Log in button	9
Figure 2. PERCEPTRON Login options	10
Figure 3. PERCEPTRON - Overview of User Interface	11
Figure 4. PERCEPTRON - Overview of Basic Parameters	12
Figure 5. PERCEPTRON - Overview of Experimental Parameters	13
Figure 6. PERCEPTRON - Overview of De Novo Sequencing Parameters	14
Figure 7. PERCEPTRON - Overview of Protein Modifications Parameters	15
Figure 8. PERCEPTRON - Overview of Scoring Component Weight	16
Figure 9. Summary Results window showing candidate proteins	17
Figure 10. Detailed Results window showing candidate proteins	17
Figure 11. Conversion of raw to mzXML	18
Figure 12. Conversion of raw to mzML	19
Figure 13. Conversion of mzXML to MGF	19
Figure 14. Load Default Parameters	20
Figure 15. Selecting Protein Database	20
Figure 16. Downloading Visual Studio Code (VS Code) setup	21
Figure 17. Choosing destination folder to download VS Code setup	21
Figure 18. Steps to install Visual Studio Code setup	22
Figure 19. Downloading SQL Server Management Studio (SSMS) setup	23
Figure 20. Steps to install SQL Server Management Studio (SSMS)	23
Figure 21. Downloading Node.js (version 8.11.1) setup	24
Figure 22. Steps to install Node.js setup	25
Figure 23. Downloading SQL 2017 setup	26
Figure 24. Steps to install SQL 2017 setup	26
Figure 25. Downloading Angular 1.7.4 setup	27
Figure 26. Steps to install Angular 1.7.4	28
Figure 27. SQL Server Management Studio Pop Up	29
Figure 28. Building PERCEPTRON API. (a) Opening solution PerceptronAPI.sln; (b) Opening WebApiConfig.cs from Solution Explorer of Visual Studio	30

Figure 29. Searching for a statement in PERCEPTRON API solution	30
Figure 30. Copying Server Name from SQL Server Management Studio	30
Figure 31. (a) Opening Web.config from Solution Explorer of Visual Studio (b) Changing data source (Server Name) in Web.config	31
Figure 32. Changing server name in Sqldatabase.cs	31
Figure 33. Building PERCEPTRON API	31
Figure 34. Steps to build PERCEPTRON API	32
Figure 35. Publishing Perceptron API	32
Figure 36. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine”	33
Figure 37. Opening PerceptronFrontEnd folder in Visual studio code	34
Figure 38. Installing extension in VS Code	35
Figure 39. Building Perceptron Front-End	35
Figure 40. Hosting Perceptron Front-End on system with default URL	36
Figure 41. Steps to build PERCEPTRON Front-End	37
Figure 42. Copying Server Name from SQL Server Management Studio	38
Figure 43. Building Perceptron Local Service. (a) Opening PerceptronLocalService.sln; (b) Opening App.config from Solution Explorer of Visual Studio	38
Figure 44. Changing data source value (Server Name) in App.config	38
Figure 45. Steps to execute Protein Search Query	39
Figure 46. Testing Perceptron Local Service	40

3. Introduction to PERCEPTRON

This chapter introduces the user to the PERCEPTRON platform along with a description of its features.

3.1. About PERCEPTRON

PERCEPTRON is a freely available web-based proteoform identification pipeline for Top-Down Proteomics (TDP). Top-down proteomics is an emerging experimental protocol for analysis of intact proteoforms. PERCEPTRON search pipeline brings together algorithms for: (i) intact mass tuning, (ii) *de novo* peptide sequence tag extraction, (iii) *in silico* spectral comparison, (iv) identification of post-translational modifications as well as truncated proteins, and (v) a novel composite scoring scheme for candidate protein scoring. PERCEPTRON achieves high performance by leveraging NVIDIA GPU technology coupled with Microsoft ASP.NET and ANGULAR frameworks. The search results obtained include a list of proteins, their scores and details on the matching information. This information can be visualized as well as downloaded. Overall, PERCEPTRON is aimed at filling the crucial void of open-source and open-architecture protein identification software for TDP data, employing state-of-the-art algorithms.

3.2. Features

The salient features of the pipeline are summarized below:

- **Graphical User Interface (GUI)** - A set of rich and intuitive graphical user interface has been developed for setting up the search parameters as well as for integrating the main components of the engine.
- **Whole Protein Molecular Weight Estimation** - The protein identification begins with the tuning of precursor protein's monoisotopic MW (MS1) as guided by its fragmentation spectra (MS2). Relative abundances and mass/charge (m/z) ratios are used to calculate the consensus MW which is then employed in the search and scoring process.
- **Peptide Sequence Tag Extractor** - Peptide sequence tag ladders (PST) are extracted from the spectra by enumerating successive peaks having MW differences equal to an amino acid and within the user specified mass tolerance. Protein database is then filtered for proteins reporting these PSTs. The length of PST ladders, cumulative mass off-sets and relative abundances are used in calculating the PST scores.
- ***In silico* fragmentation** – *In silico* fragments of candidate proteins are generated by the user selected fragmentation techniques. *In vitro* and *in silico* spectral comparisons are performed and scored.
- **Post-translational Modification (PTM) Search** - Support for predicting typical PTMs has been provided in the tool. Users can select and search variable and fixed PTMs of their choice along with blind-PTMs by simply selecting them from the GUI.
- **Multifactorial Composite Scoring System** - A multifactorial candidate protein scoring scheme incorporating the aforementioned algorithms has been developed. User customization of the parameters and weights in the scoring function is admitted via a GUI.
- **Single and Batch Search** – PERCEPTRON provides support for search in single as well as batch modes. Towards an automated processing of multiple spectral data files, a batch processing mode allows for the selection of multiple files from the folder by clicking the attach file button. The

PERCEPTRON: User Manual

experimental spectra, search parameters and results are automatically stored in the project directory for further processing and visualization.

4. Hardware and Software

4.1. Hardware

For in-house deployment of PERCEPTRON, the hardware requirement includes a graphical processing unit (GPU) that supports CUDA TOOLKIT 7.0.

4.2. Software

Compilation and software build of PERCEPTRON requires the following software:

- Windows Server 2012 R2
- Visual Studio 2013
- Angular 1.7.4
- Node.js 8.11.1
- SQL Server Management Studio (17.6 or 17.9.1)
- CUDA TOOLKIT 7.0
- CUDAFy.NET.1.29.5576.13786
- Microsoft Windows Server v6.2
- .NET Framework 4.5

4.3. Testing

The freely available public deployment of PERCEPTRON is hosted on a Dell Power Edge R730, 2 x Intel Xeon E5-2620, 160 GB RAM (16GBx10) and an NVIDIA Tesla K40C (2880 Cores). Following Windows versions are compatible with PERCEPTRON:

- Windows 8.1
- Windows 7
- Win Server 2012 R2
- Win Server 2008 R2

5. Video Tutorials

Several video tutorials have been developed for employment of PERCEPTRON in protein search. These tutorials are available as a playlist at: https://www.youtube.com/playlist?list=PLaNVq-kFOn0Z_7b-iL59M_CeV06JxEXmA

6. Getting Started with PERCEPTRON

PERCEPTRON manual, samples and issues database is freely available (under the MIT open license) at (<https://perceptron.lums.edu.pk/>)

To initiate use of PERCEPTRON, the user has to register and log in by clicking on the link provided above.

The screenshot displays the PERCEPTRON homepage. At the top right, there is a 'Log in' button highlighted with a red rectangle. The left sidebar contains navigation links: About Perceptron, Getting Started, Protein Search Query (highlighted), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Acknowledgment, and Contact. The main content area features the title 'PERCEPTRON - Proteoform Identification Pipeline for Top-Down Proteomics' and a welcome message. Below the text, a detailed flowchart illustrates the 'Protein Search Pipeline'. The pipeline starts with 'Input Data Preparation' (Protein Sample, Protein Separation, Ionization, Mass Spectrometer, MS Spectra) and proceeds through 'Protein Search Pipeline' steps: de novo PST(s) Extraction, Filter Candidate Proteins using PST(s), Calculating Final Score, and Final Protein List. It also includes 'Filter Protein DB on MW', 'In Silico Fragmentation', 'Post Translational Modifications', and 'Spectral Comparison'. The diagram uses various icons and text boxes to represent these steps and the data flow.

Figure 1. PERCEPTRON Homepage and Log in button

Click the '**Log in**' button on the top right. A window will appear. Enter user credentials to proceed.

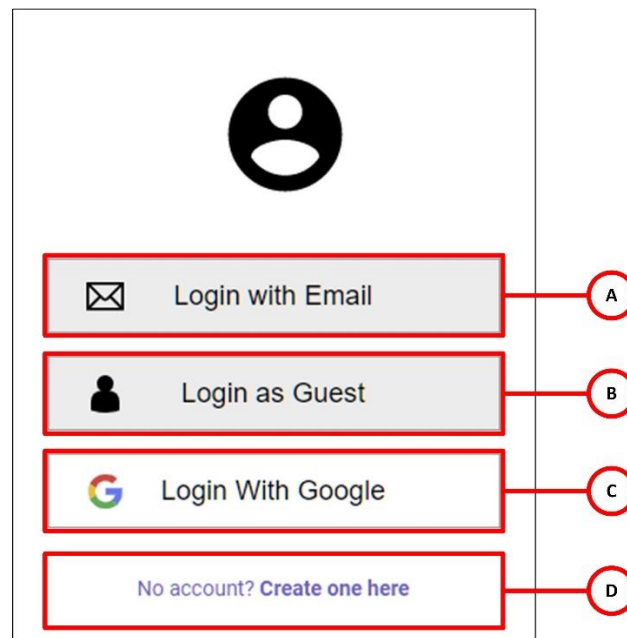


Figure 2. PERCEPTRON Login options

- A. User can login with an existing email account
- B. Enables user to login as a guest, perform search in PERCEPTRON. Note that results will not be saved in this case
- C. User can also use their Google account to login
- D. Otherwise, users can create a new local account in order to login

7. GUI Description

This chapter presents the interface overview for user facilitation.

7.1 Window 1: PERCEPTRON Tool for Top-down Proteomics

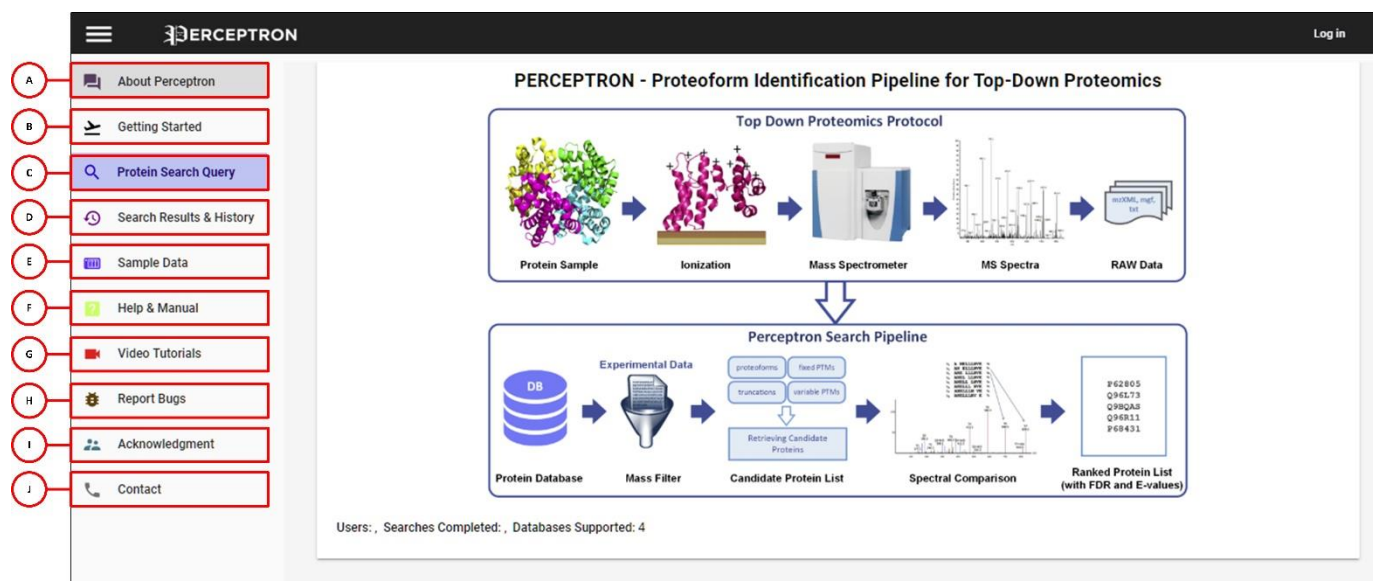


Figure 3. PERCEPTRON - Overview of User Interface

- A. About PERCEPTRON: A next-generation top-down proteoform search and identification platform
- B. Getting Started: Quick guide to proteoform search and identification using PERCEPTRON
- C. Protein Search Query: Job submission - Search top-down proteomics data files to identify and characterize proteoforms
- D. Search Results & History: View search results and history
- E. Sample Data: Sample top-down mass spectrometry data for use with PERCEPTRON
- F. Help & Manual: Get assistance with using PERCEPTRON and download manual
- G. Video Tutorials: View step-by-step video tutorials demonstrating usage of PERCEPTRON
- H. Report Bugs: Report problems and issues here
- I. Acknowledgement: PERCEPTRON project team members
- J. Contact: Contact us for further information

In order to start protein search, click on the 'Protein Search Query' tab and the following window will appear:

7.2 Window 2: Protein Search Query

7.2.1 Basic Parameters

The screenshot shows the 'Submit Protein Identification Query' interface. On the left is a sidebar with navigation links: About Perceptron, Getting Started, Protein Search Query (highlighted), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Acknowledgment, and Contact. The main content area is titled 'Submit Protein Identification Query' and contains two main sections. The 'Basic Parameters' section has a 'Load Default Parameters' link and several input fields: 'Protein Search Title' (labeled A), 'Intact Protein Mass' (labeled B), 'Upload Mass-Spectrometry Data File(s):' with a file selection button (labeled D), 'Select Database' (labeled C), 'Enter your E-mail Address *' (labeled E), and 'Recieve Top' with a dropdown menu (labeled F). Below this is the 'Set Experimental Parameters' section, which includes expandable sections for 'Set Experimental Parameters', 'Set De Novo Sequencing Parameters', 'Set Protein Modifications Parameters', and 'Set Scoring Components Weight'. At the bottom right are 'Reset' and 'Submit' buttons. The footer contains 'Contact Us' and '(c) 2019 Biomedical Informatics Research Laboratory (BIRL)'.

Figure 4. PERCEPTRON - Overview of Basic Parameters

- A. In order to start protein search, user must enter the 'Protein Search Title' (for example: "MyProject")
- B. Enter intact protein mass (MS1)
- C. Select protein database
- D. Browse and upload experimental data (.mzXML/ .MGF/ .txt) for Single mode; Peak-list files for Batch mode
- E. Enter the email address at which the link for downloading results will be sent. Note that users are notified about the availability of search results upon completion of the job
- F. Provide the number of candidate protein hits to be received in results

7.2.2 Set Experimental Parameters

The screenshot displays the PERCEPTRON web application interface. The sidebar on the left contains navigation links: About Perceptron, Getting Started, Protein Search Query (highlighted), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Acknowledgment, and Contact. The main content area is titled 'Submit Protein Identification Query'. It includes a 'Basic Parameters' section with fields for Protein Search Title, Intact Protein Mass (Da), Upload Mass-Spectrometry Data File(s) (0 Files Selected), Select Database, Enter your Email Address *, Recieve Top, and Results. Below this is the 'Set Experimental Parameters' section, which is expanded. It contains several input fields and checkboxes: Mass Mode (radio buttons for MH+ and M(Neutral)), Intact Mass Tolerance (Da), Peptide Tolerance (ppm), Filter Database using MS1 (checkbox), Tune Intact Protein Mass (checkbox), Handle Truncated Proteoforms (checkbox), Neutral Mass Loss (Da), Select Fragmentation Type (dropdown), and Select Corresponding Special Ions (dropdown). At the bottom of the main content area are sections for Set De Novo Sequencing Parameters, Set Protein Modifications Parameters, and Set Scoring Components Weight. A 'Reset' button and a 'Submit' button are located at the bottom right of the interface.

Figure 5. PERCEPTRON - Overview of Experimental Parameters

- Select Mass Mode. MS data can only be provided in either m/z form with $z = 1$ or neutral masses
- Set the tolerance value for Protein Mass and select its unit
- Select the tolerance value for Peptide and select its unit
- Select the 'Fragmentation type' from drop down menu
- User can filter database by checking the option 'Filter Database using MS1'
- Check 'Handle Truncated Proteoforms' to allow search for truncated proteoforms
- Provide the value of Neutral loss, if any
- Choose the corresponding special ions for the type of fragmentation selected (i.e. a', b', y', z'', a*, b*, y*, z' ions)
- Check the option 'Tune Intact Protein Mass' to allow for tuning of MS1 using MS2 data

7.2.3 Set *De Novo* Sequencing Parameters

The screenshot shows the PERCEPTRON web application interface. The left sidebar contains navigation links: About Perceptron, Getting Started, Protein Search Query (selected), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Acknowledgment, and Contact. The main content area is titled 'Submit Protein Identification Query'. It includes a 'Basic Parameters' section with fields for 'Protein Search Title', 'Intact Protein Mass', 'Da', 'Upload Mass-Spectrometry Data File(s)', '0 Files Selected', 'Select Database', 'Enter your E-mail Address *', 'Recieve Top', and 'Results'. Below this is the 'Set Experimental Parameters' section. The 'Set *De Novo* Sequencing Parameters' section is expanded and contains the following parameters:

- A:** ☐ Enable PST Filtering
- B:** Tolerance For Each Hop: Da
- C:** Minimum Tag Length: [dropdown menu]
- D:** Maximum Tag Length: [dropdown menu]
- E:** Set the overall mass error tolerance for th... [text input]

Below these are sections for 'Set Protein Modifications Parameters' and 'Set Scoring Components Weight'. At the bottom right are 'Reset' and 'Submit' buttons.

Figure 6. PERCEPTRON - Overview of *De Novo* Sequencing Parameters

- Check 'Enable PST Filtering' to filter PSTs
- Set the 'Tolerance for each Hop'
- Tags will be filtered above the minimum length of PST selected from the drop down menu by the user
- Tags will be filtered below the maximum length of PST selected from the drop down menu by the user
- Overall mass error tolerance shows error margin for the whole PST

7.2.4 Set Protein Modifications Parameters

The screenshot shows the PERCEPTRON web interface. The sidebar on the left contains links: About Perceptron, Getting Started, Protein Search Query (selected), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Acknowledgment, and Contact. The main content area is titled 'Submit Protein Identification Query'. It includes a 'Basic Parameters' section with fields for 'Protein Search Title', 'Intact Protein Mass', 'Da', 'Upload Mass-Spectrometry Data File(s)', 'Select Database', and 'Enter your E-mail Address *'. Below this are expandable sections: 'Set Experimental Parameters', 'Set De Novo Sequencing Parameters', and 'Set Protein Modifications Parameters' (which is expanded). The 'Set Protein Modifications Parameters' section contains:

- Terminal Modification** (A): A dropdown menu.
- Blind-PTM Search** (B): A checkbox.
- PTM Tolerance** (C): A dropdown menu.
- Methionine Chemical Modifications** (D): A dropdown menu.
- Cysteine Chemical Modifications** (F): A dropdown menu.
- Fixed Modifications** (G): A dropdown menu.
- List of Modifications**: A scrollable list on the left with items: Acetylation_A, Acetylation_K, Acetylation_M, Acetylation_S, Amidation_F, Hydroxylation_p, Methylation_K, Methylation_R, N-Linked Glycosylation_N, and O-Linked Glycosylation_T. Arrows indicate selection options.

 At the bottom are 'Reset' and 'Submit' buttons.

Figure 7. PERCEPTRON - Overview of Protein Modifications Parameters

- Allows the user to select specified terminal modifications. PERCEPTRON handles four cases: 1) None – No modification, 2) NME – N terminal methionine excision, 3) NME_ACETYLTATION – N terminal acetylation with initiator methionine removed, and 4) M_ACETYLTATION – N terminal methionine acetylation
- Select whether to perform Blind-PTM search and unknown modifications
- Set the tolerance value for Post Translational Modification (PTM) site selection
- Allows the user to select instrument specific modification on Methionine
- User can opt for required Variable 'Post translation Modifications' from the list of modifications
- Allows the user to select instrument specific modification on Cysteine
- Similarly, various 'Fixed Modifications' are also selected from the list

7.2.5 Set Scoring Components Weight

The screenshot displays the PERCEPTRON web application interface. On the left is a sidebar with navigation links: About Perceptron, Getting Started, Protein Search Query (highlighted), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Acknowledgment, and Contact. The main content area is titled 'Submit Protein Identification Query'. It contains several sections: 'Basic Parameters' with fields for Protein Search Title, Intact Protein Mass, and a dropdown for Select Database; 'Upload Mass-Spectrometry Data File(s):' with a file upload button showing '0 Files Selected'; and an 'Enter your E-mail Address *' field. Below these are expandable sections for 'Set Experimental Parameters', 'Set De Novo Sequencing Parameters', 'Set Protein Modifications Parameters', and 'Set Scoring Components Weight'. The 'Set Scoring Components Weight' section is expanded and contains three sliders: 'Intact Protein Mass Score Weightage (%)', 'Peptide Sequence Tags Score Weightage (%)', and 'Spectral Comparisons Score Weightage (%)'. A red rectangular box highlights these three sliders. A red circle with the letter 'A' points to the 'Spectral Comparisons Score Weightage (%)' slider. At the bottom right of the main content area are 'Reset' and 'Submit' buttons.

Figure 8. PERCEPTRON - Overview of Scoring Component Weight

- A. Set the respective weights of the Scoring Components from by shifting the slider left or right accordingly

7.3 Window 3: Summary and Detailed Results View

Summary Results View

Filter

Protein Rank	Protein Name	Protein ID	Molecular Weight	Terminal Modification	No. of Modification(s)	Protein Score
1	Protein 0	Q5TBE3	11359.7861	no	2	0.3159
2	Protein 1	P62805	11473.4114	no	2	0.0340
3	Protein 2	Q99525	11115.1934	no	2	0.0274
4	Protein 3	A6NFX4	11363.5022	no	2	0.0240
5	Protein 4	Q9C005	11355.8325	no	2	0.0227
6	Protein 5	Q9P1G2	11483.6480	no	2	0.0204
7	Protein 6	Q5T752	11334.5254	no	2	0.0183
8	Protein 7	Q9BTM9	11486.0357	no	2	0.0141
9	Protein 8	Q5T292	11431.1326	no	2	0.0136
10	Protein 9	P09341	11407.2531	no	2	0.0129
11	Protein 10	P19875	11494.3653	no	2	0.0129

Items per page: 5 0 of 0

* Click on the table row for corresponding detailed results.

Contact Us (c) 2018 Biomedical Informatics Research Laboratory (BIRL)

Figure 9. Summary Results window showing candidate proteins

Proteins found after the search are reported along with their protein ID, molecular weight and score. Users can click on any protein to see the ‘Detailed Result View’ of the protein. Click on the ‘protein ID’ to go to the detailed UniProt view of the protein.

Detailed Result View

User Search Parameters

General Results

Protein Rank:	2	Protein ID:	P62805	Protein Name:	P62805
Protein Score:	0.034039	Molecular Weight:	11473.411430	# Matched Fragments:	12
Terminal Modification:	NME	Truncation:	No	# Modifications:	3

* Click on the Protein ID to access its information.

Protein Search Time (in Seconds)

Total:	00:00:00.0470323	Molecular Weight Module:	N/A	Peptide Sequence Tag Module:	00:00:00.0111953
Spectral Comparison Module:	00:00:00.0000004	Post-Translational Modification Module:	00:00:00.0039530	Truncation Module:	00:00:00.0470323

1. MSGRGKGGKG 2. LGKGAKRHR 3. KVLRDNIQGI 4. TKPAIRRLAR 5. RGGVKRISGL 6. IYEETRGVLK 7. VFLENVIRDA 8. VITYTEHAKRK 9. TVTAMDVVYA 10. LKRGQRTLYG 11. FGG

Contact Us (c) 2018 Biomedical Informatics Research Laboratory (BIRL)

Figure 10. Detailed Results window showing candidate proteins

8. Search

8.1 File Formats

PERCEPTRON provides support for plain text files (data in columns containing mass to charge ratios (m/z) and relative intensities), eXtensible Markup Language (XML) files with m/z and relative abundances (mzXML)¹, Mass Spectrometry Markup Language (mzML)^{2,3} and Mascot Generic Format (MGF)⁴ data formats in both single and batch file processing modes.

8.1.1 Raw to mzXML File Format Conversion

User can convert raw data files to mzXML file format by using MS-Convert⁵.

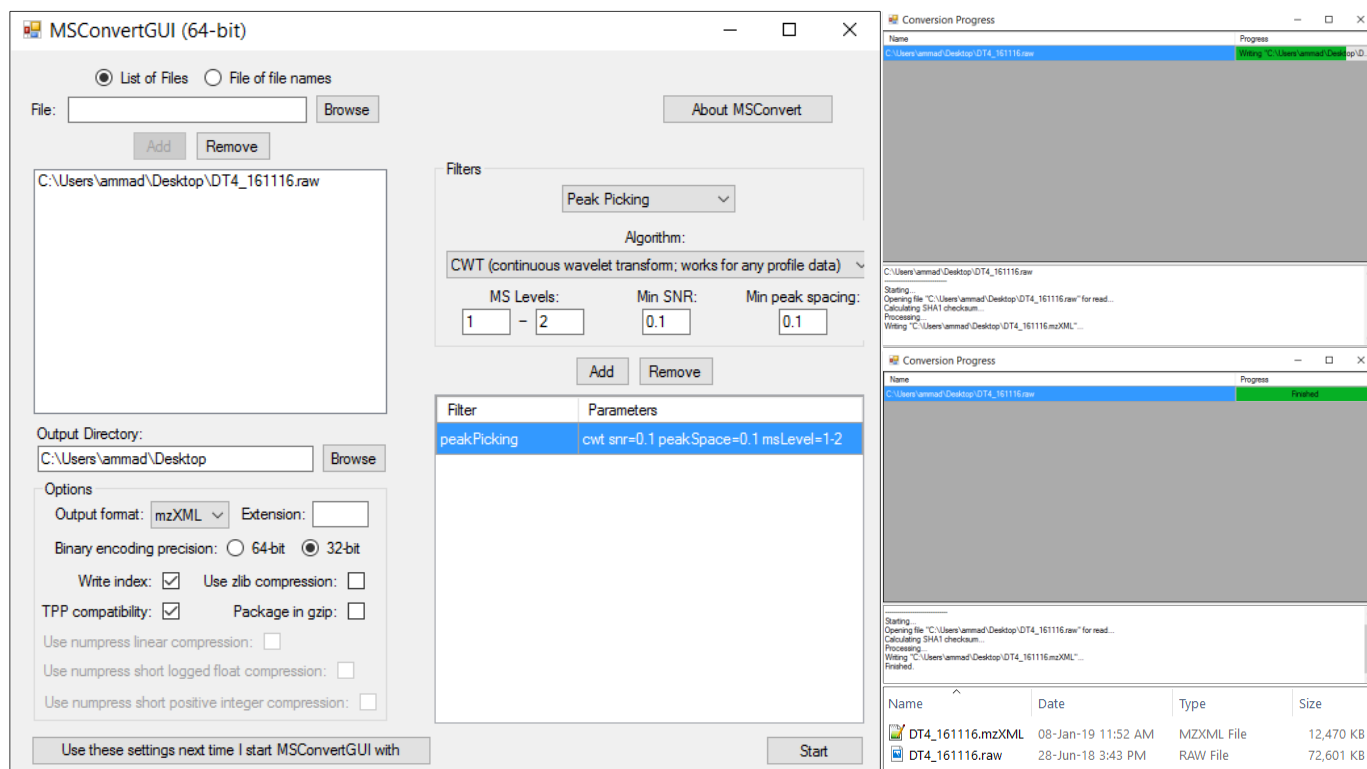


Figure 11. Conversion of raw to mzXML

8.1.2 Raw to mzML File Format Conversion

Raw data files can be converted to mzML file format by using MS-Convert⁵.

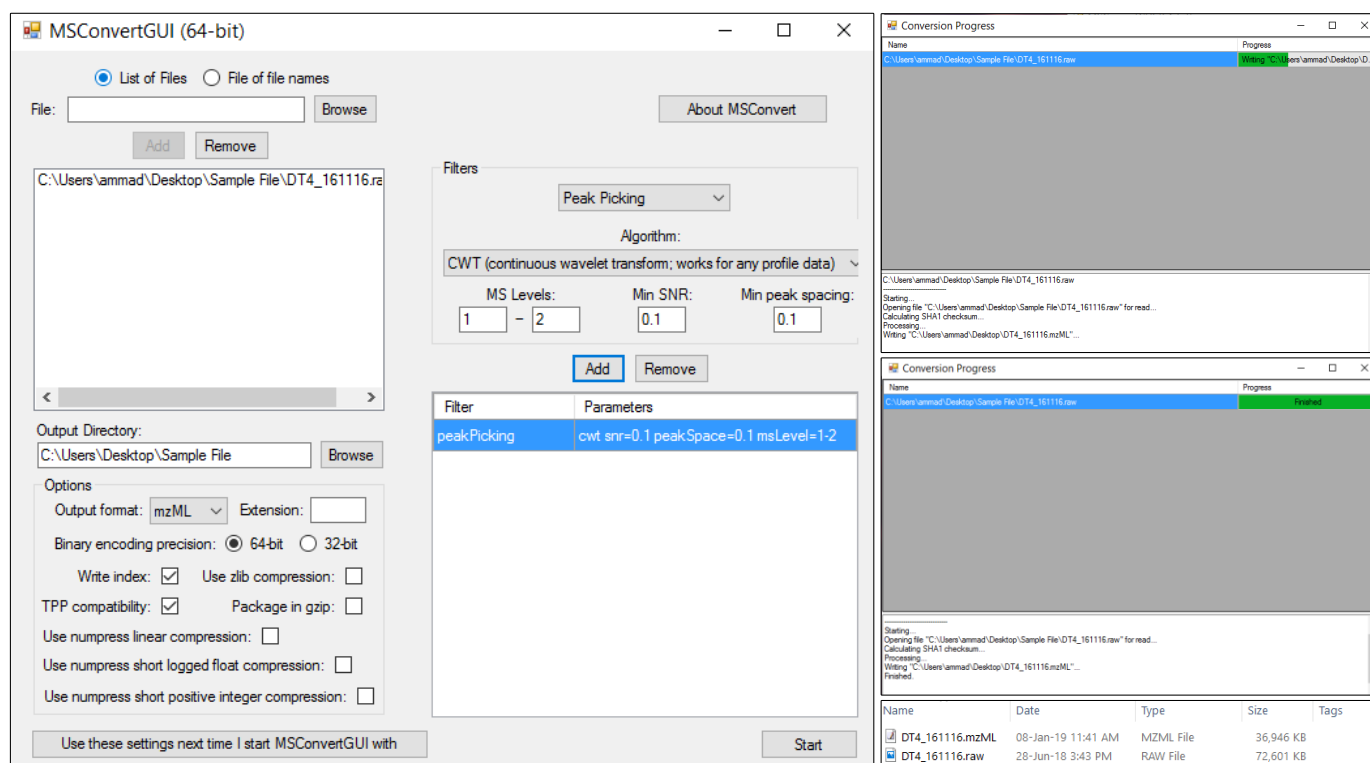


Figure 12. Conversion of raw to mzML

8.1.3 MzXML to MGF File Format Conversion

User can convert mzXML files to MGF using MS-Decov⁶.

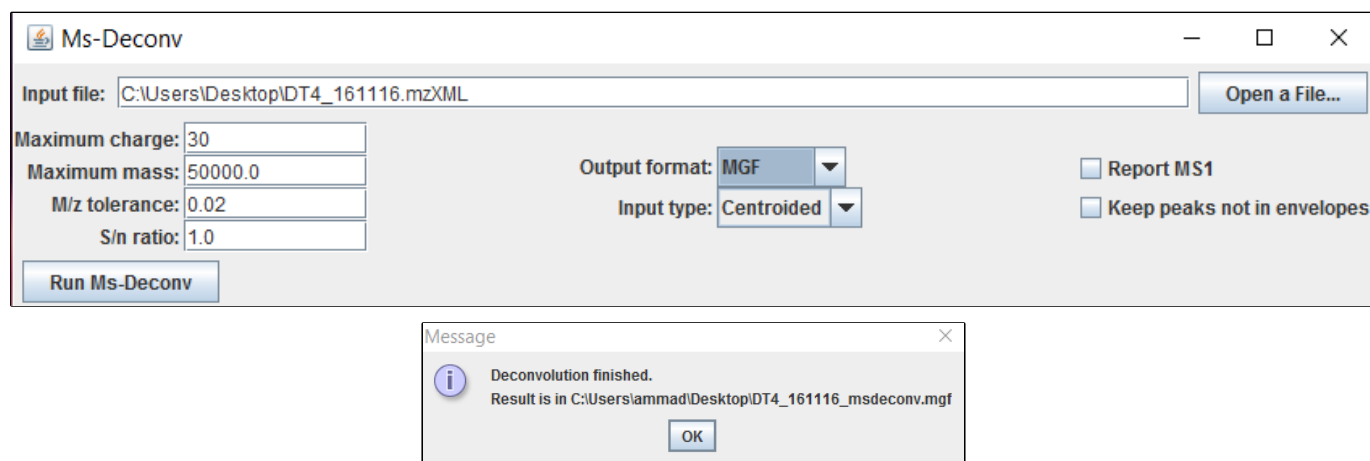


Figure 13. Conversion of mzXML to MGF

8.1.4 MGF to Flat Text File Format Conversion

PERCEPTRON converts MGF files to flat text (peak list) using built-in custom file reader.

8.2 Parameters

PERCEPTRON can employ from the following set of search parameters including: (i) Default Parameters, and (ii) Selected parameters.

How to load default Parameters?

To submit the job using default parameters, select 'Load Default Parameters' option in front of Basic Parameters.

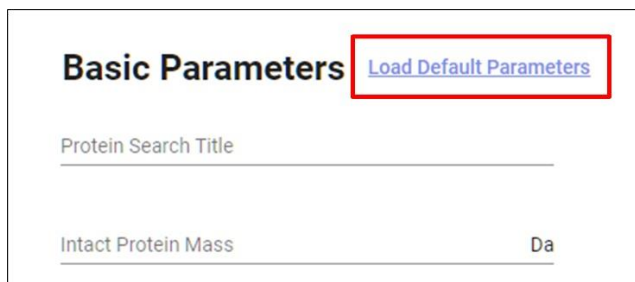
A screenshot of a web form titled "Basic Parameters". The form contains two input fields: "Protein Search Title" and "Intact Protein Mass" (with a unit "Da" to its right). A red rectangular box highlights a blue hyperlink labeled "Load Default Parameters" located to the right of the "Basic Parameters" title.

Figure 14. Load Default Parameters

8.3 Databases

SwissProt database is included in PERCEPTRON by default. User can take any protein sequence from other databases (such as Uniprot) in .fasta format.

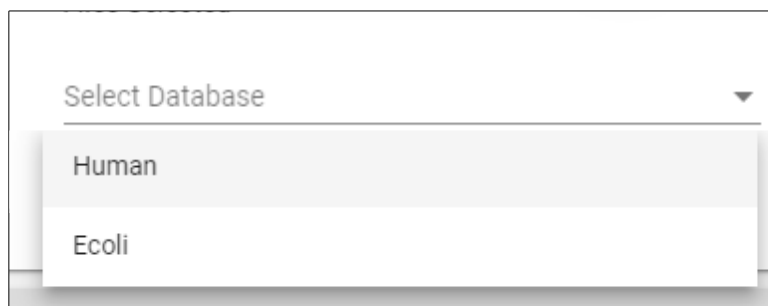
A screenshot of a web form showing a dropdown menu labeled "Select Database". The dropdown is open, displaying two options: "Human" and "Ecoli". The "Human" option is highlighted with a light gray background.

Figure 15. Selecting Protein Database

8.4 Modes

The search modes are auto-selected based on the number of files that the user inputs. If one file is given as input, PERCEPTRON runs single search mode whereas if multiple files are given as input, the mode switches to batch mode.

- (i) Single Search Mode
- (ii) Batch Mode

Batch mode takes more processing time as it deals with larger data. The experimental spectra, search parameters and results are automatically stored in the project directory for further processing and visualization.

9. Deploying PERCEPTRON

9.1 Installing Visual Studio Code (VS Code)

Users can download VS Code setup using the following link:

<https://code.visualstudio.com/download>

Click **'64 bit'** in front of **User Installer** to download the setup. Run the downloaded .exe file.

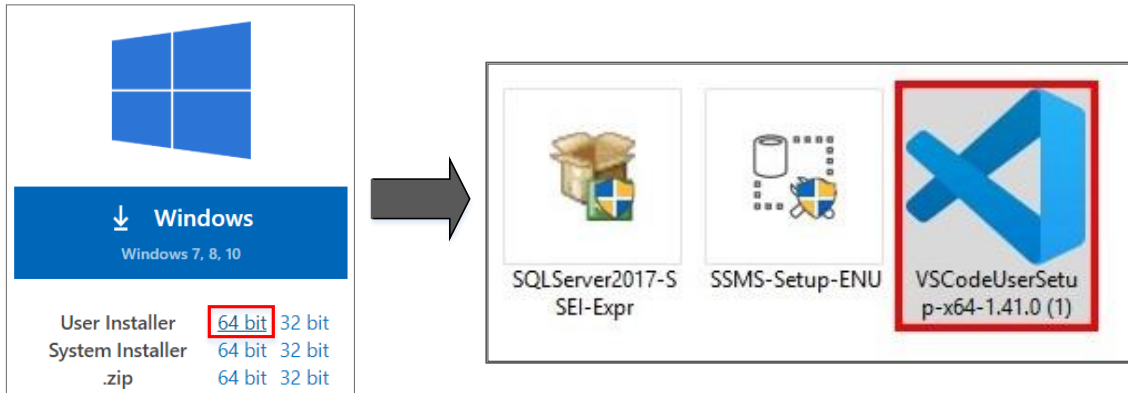


Figure 16. Downloading Visual Studio Code (VS Code) setup

Accept license terms and choose a destination folder to download Visual Studio Code setup.

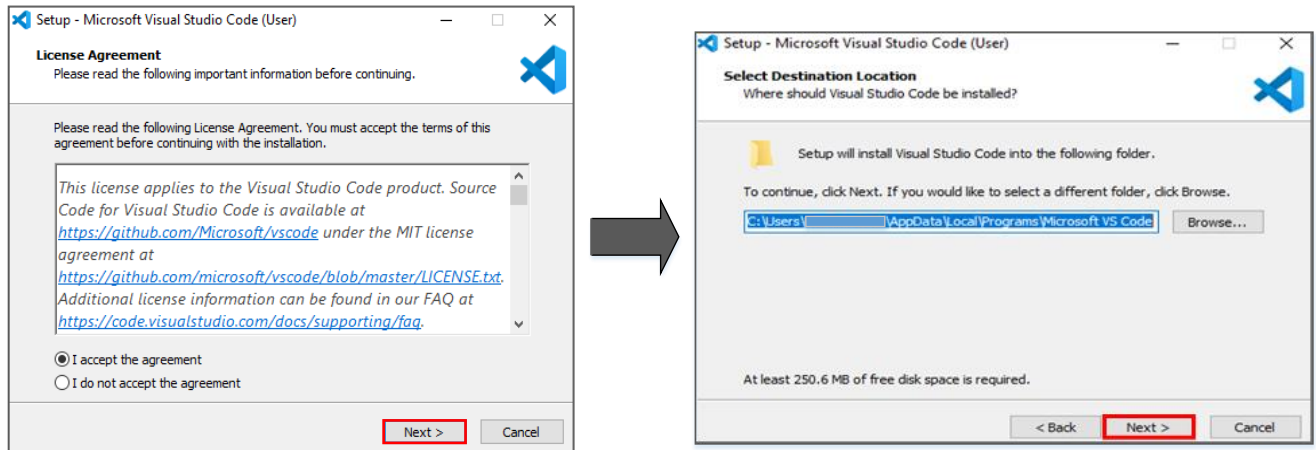
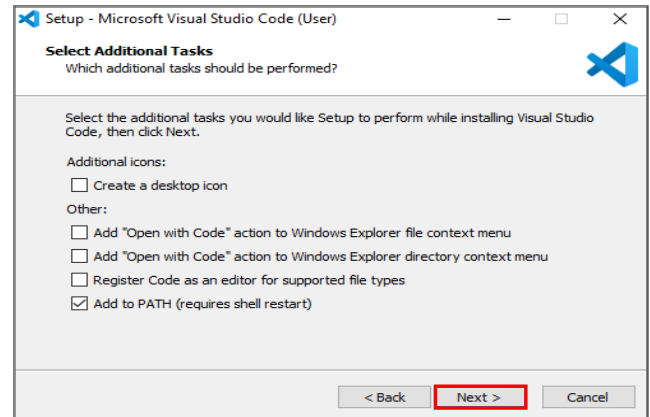
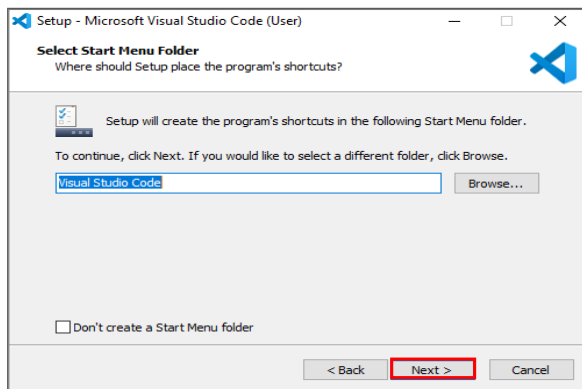


Figure 17. Choosing destination folder to download VS Code setup

PERCEPTRON: User Manual

Click **'Next'** then select **'Add to PATH.'** Click **'Next'** to proceed.



Click **'Install'** to begin installation. Click **'Finish'** to complete installation.

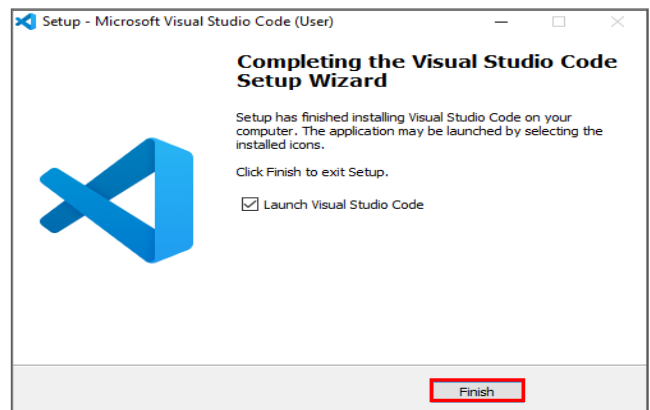
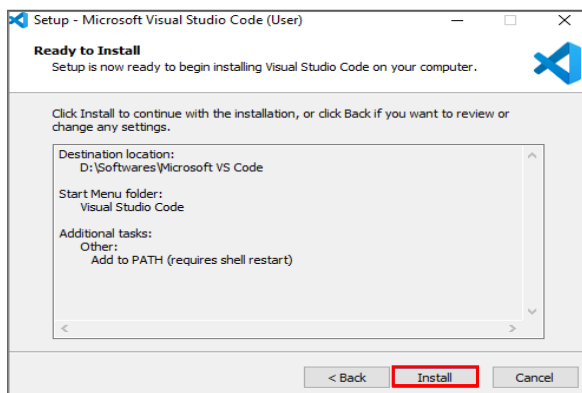


Figure 18. Steps to install Visual Studio Code setup

9.2 Installing SQL Server Management Studio (SSMS)

Users can download SSMS setup (version 17.6 or version 17.9.1) using the following link:

<https://docs.microsoft.com/en-us/sql/ssms/release-notes-ssms?view=sql-server-ver15#additional-downloads>

Click **'Download'** to download the setup. Run the downloaded .exe file.

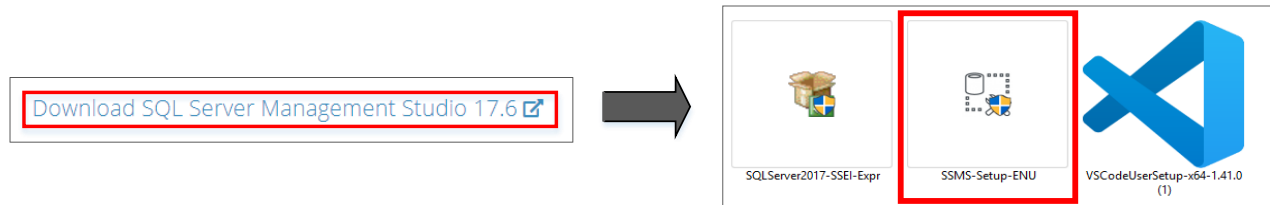


Figure 19. Downloading SQL Server Management Studio (SSMS) setup

Click **'Install'** to begin installation. Click **'Close'** to complete installation.

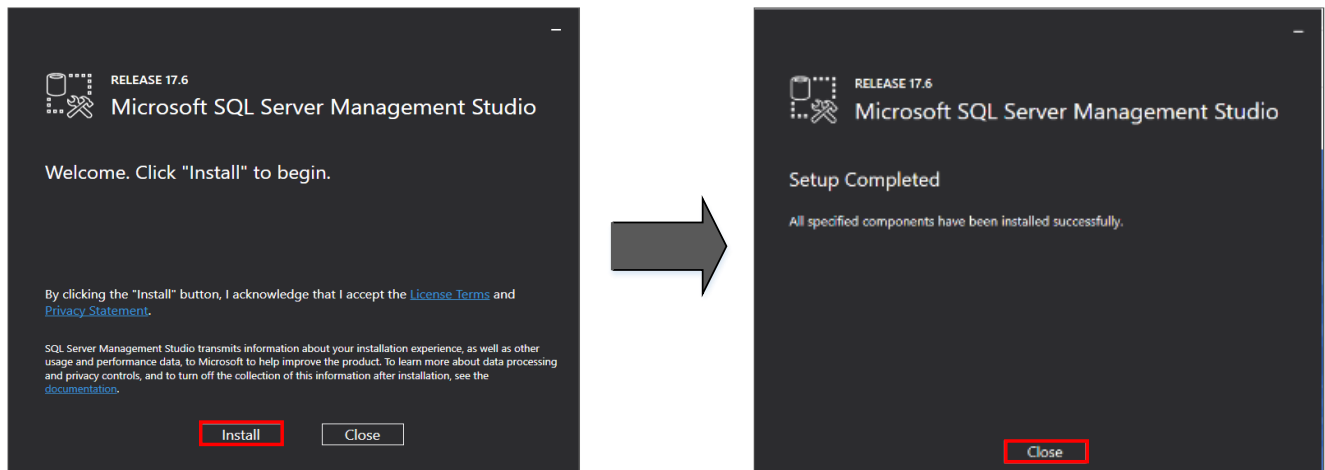


Figure 20. Steps to install SQL Server Management Studio (SSMS)

9.3 Installing Node.js.8.11.1

Users can download the setup using the following link:

<http://www.npackd.org/p/org.nodejs.NodeJS64/8.11.1>

Click on **'Download Node.js 64 bit 8.11.1'** to download node.js setup. Click on **'Run'** to proceed with downloading the .exe file.

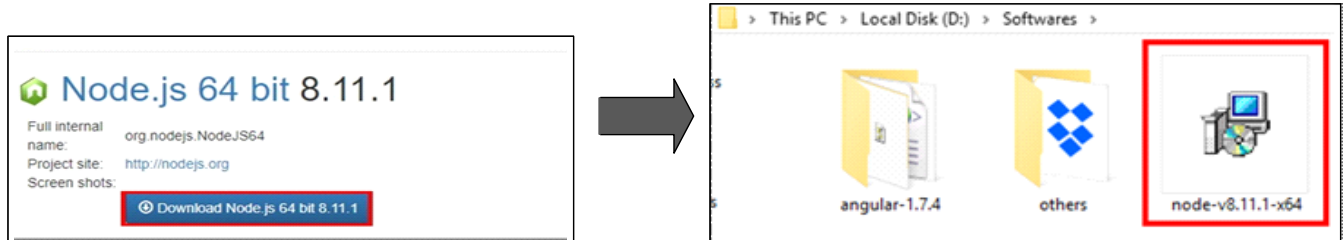
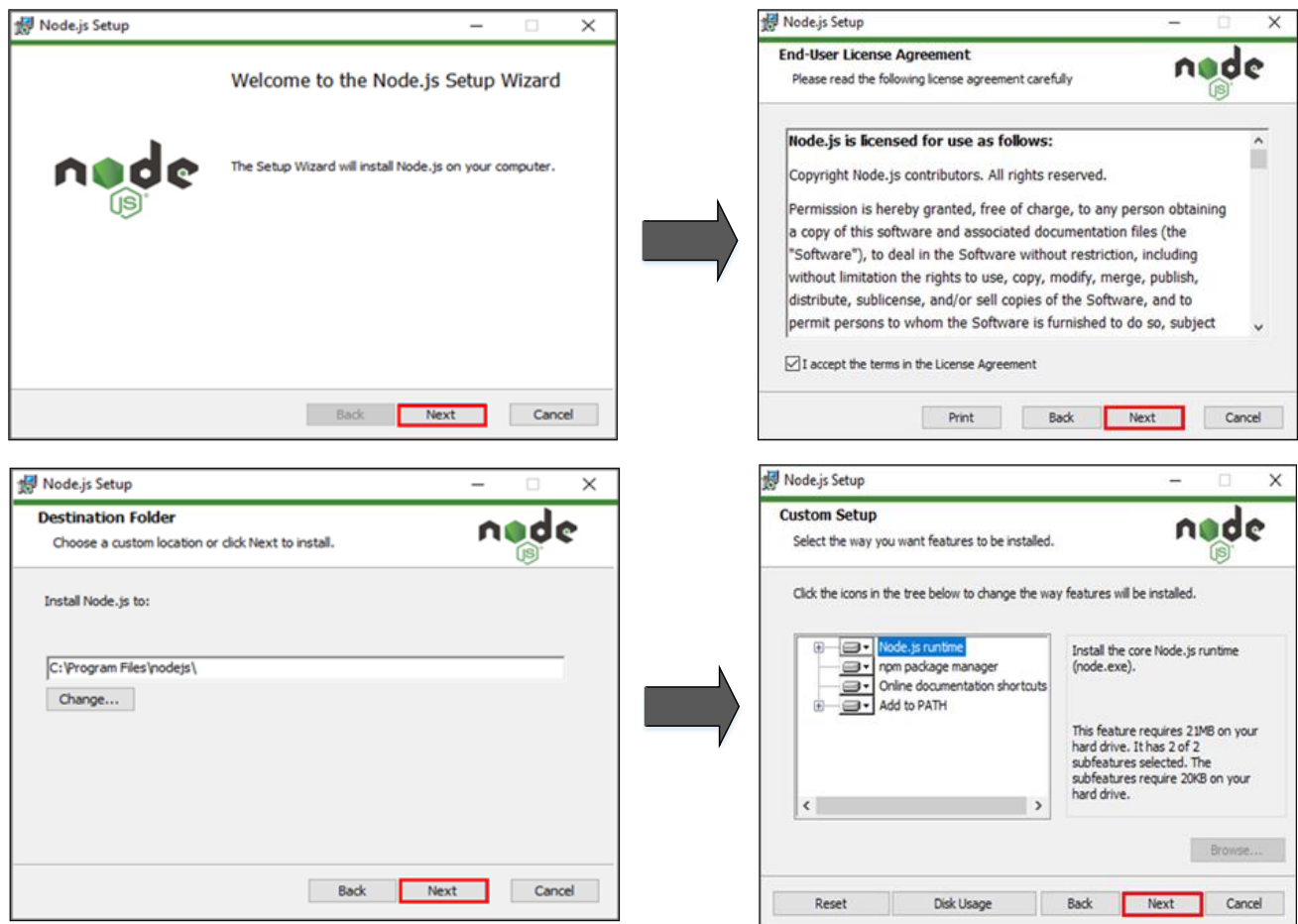
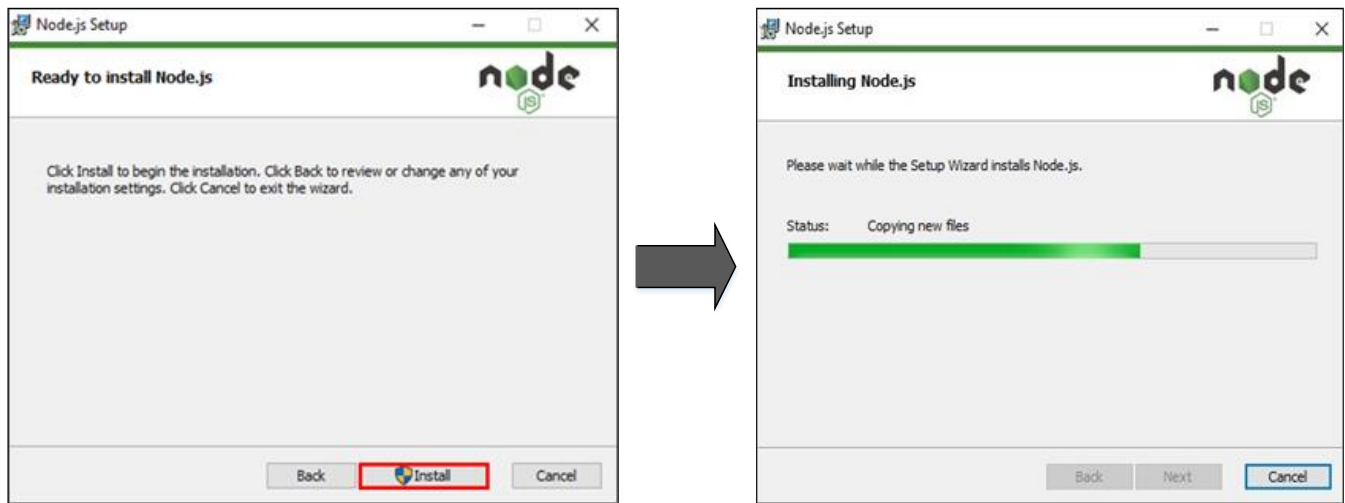


Figure 21. Downloading Node.js (version 8.11.1) setup

Click **'Next'** and accept license terms to proceed.



Click **'Install'** to begin installation.



Click **'Finish'** to complete installation.

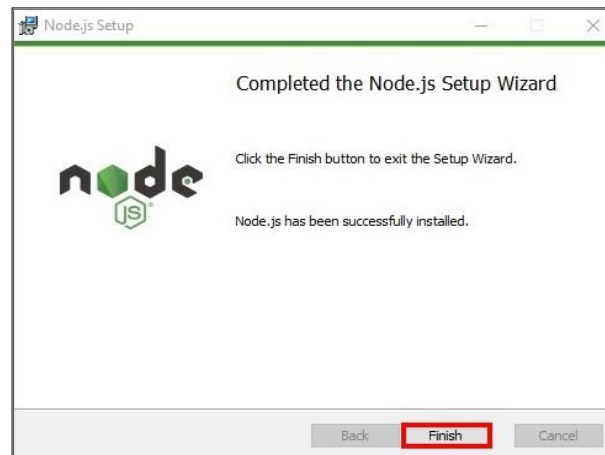


Figure 22. Steps to install Node.js setup

9.4 Installing SQL 2017

Users can download Structured Query Language (SQL) 2017 setup using the following link:

<https://www.microsoft.com/en-pk/download/details.aspx?id=55994>

Select a language and click **‘Download.’**

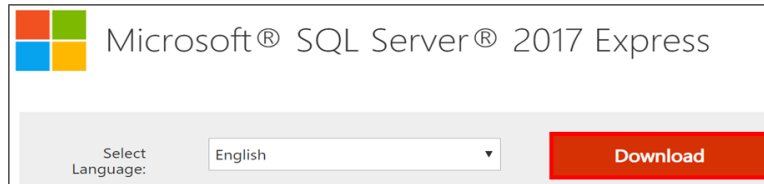
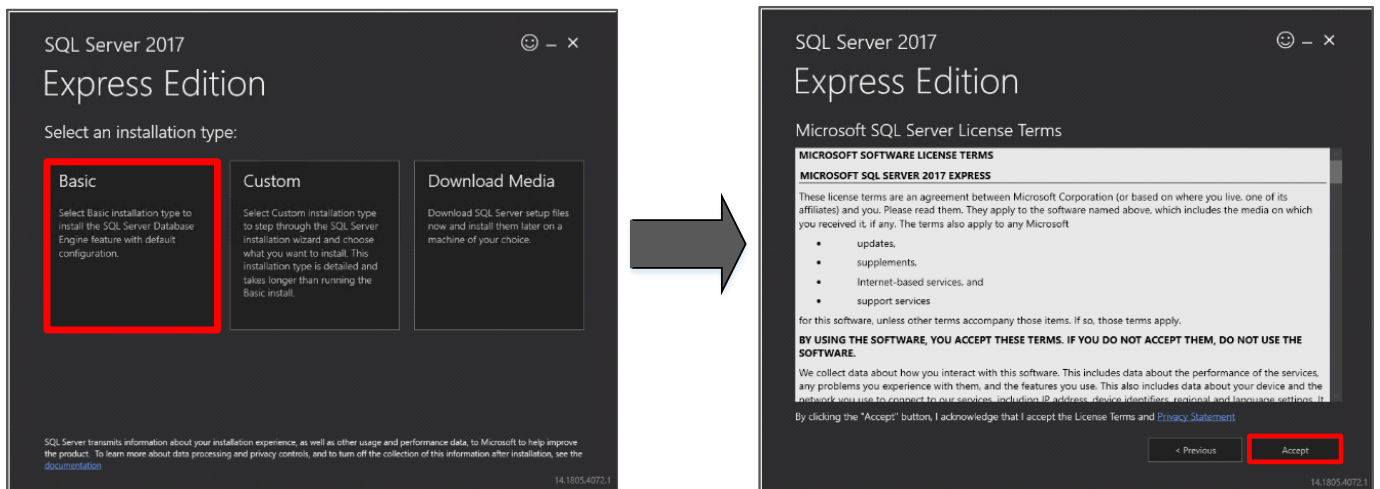


Figure 23. Downloading SQL 2017 setup

Select **‘Basic’** and then click **‘Accept’** to accept license terms.



Allow installation to proceed. Click **‘Close’** to complete installation.

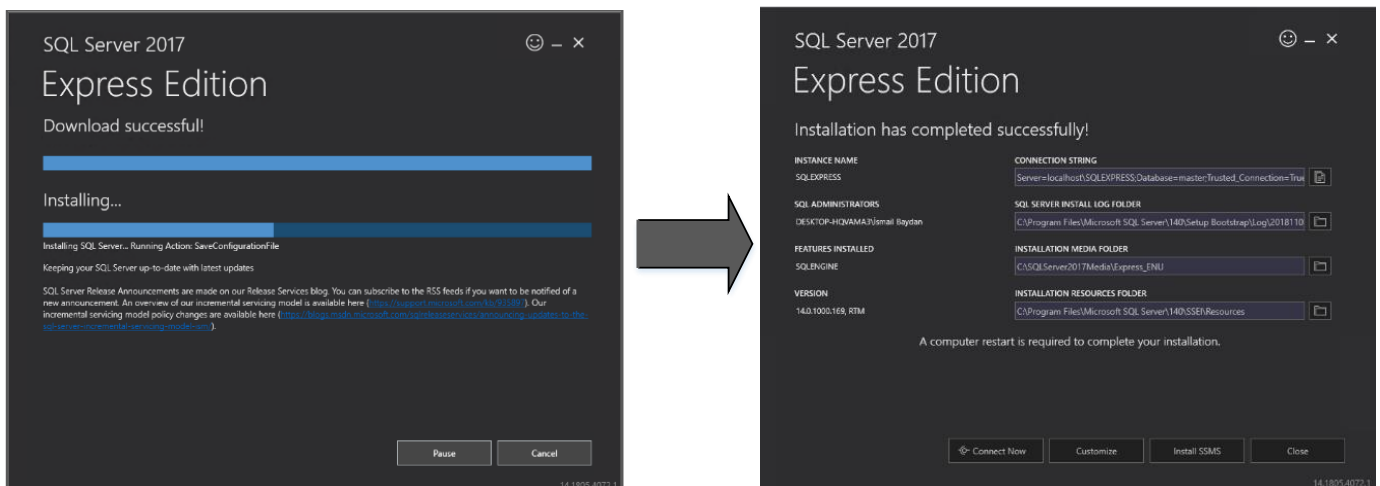


Figure 24. Steps to install SQL 2017 setup

9.5 Installing Angular 1.7.4

Users can download Angular 1.7.4 as a zip file using the following link:

<https://code.angularjs.org/1.7.4/>

Open the link and select the .zip file for installing Angular 1.7.4. Extract the contents in the file.

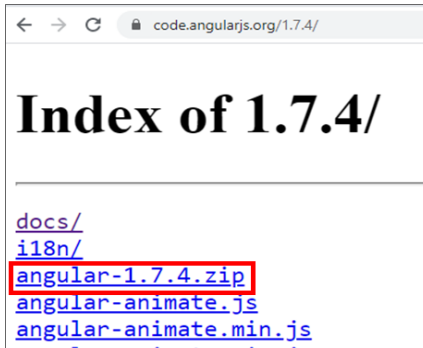
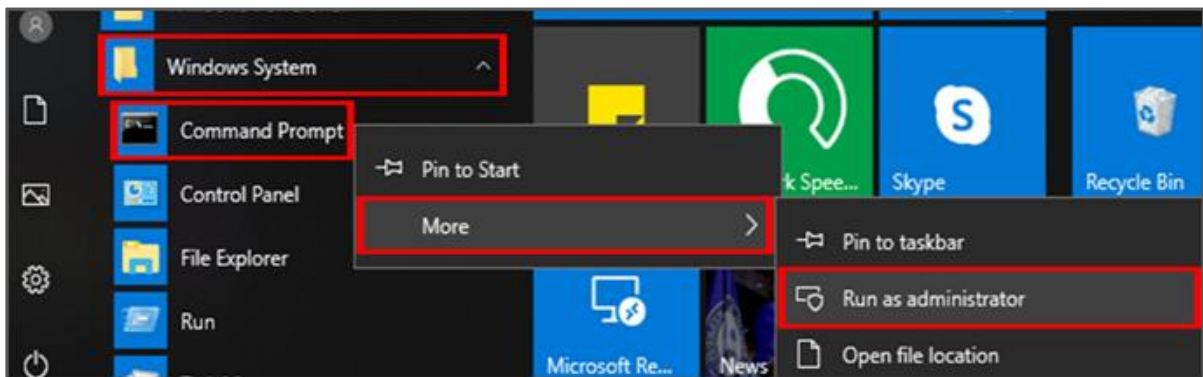


Figure 25. Downloading Angular 1.7.4 setup

Next, open Command Window by selecting **‘Windows System.’** Right click on **‘Command Prompt,’** then select **‘More’** and click **‘Run as administrator.’**



Type **'npm uninstall -g @angular/cli'** in Command Window and press **ENTER**. Next, type **'npm cache verify'** and press **ENTER** again.

```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.17134.407]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>npm uninstall -g @angular/cli
up to date in 0.031s

C:\WINDOWS\system32>
```



```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.17134.407]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>npm uninstall -g @angular/cli
up to date in 0.031s

C:\WINDOWS\system32>npm cache verify
Cache verified and compressed (~\AppData\Roaming\npm-cache\_cacache):
Content verified: 2162 (80052308 bytes)
Index entries: 3387
Finished in 6.055s

C:\WINDOWS\system32>
```

Type **'npm install -g @angular/cli@1.7.4'** and press **ENTER** to install. Type **'ng -v'** and press **ENTER** to ensure successful installation.

PERCEPTRON: User Manual

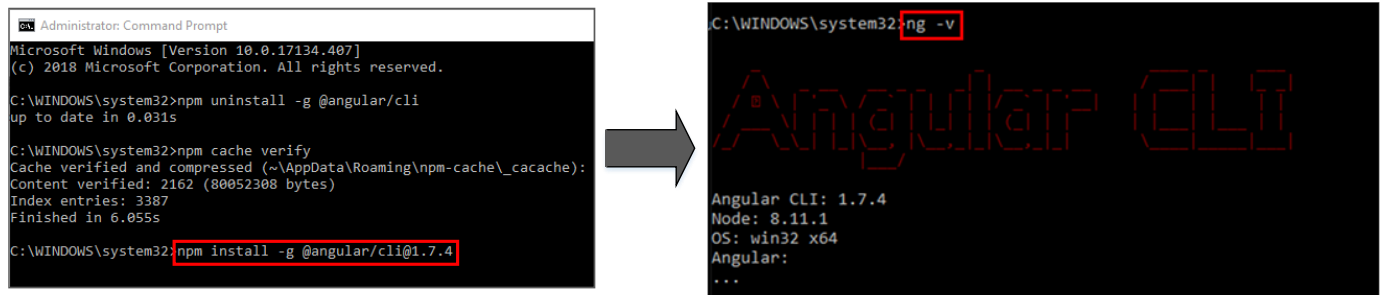


Figure 26. Steps to install Angular 1.7.4

9.6 Importing files into SQL

First, open SQL Server Management Studio.

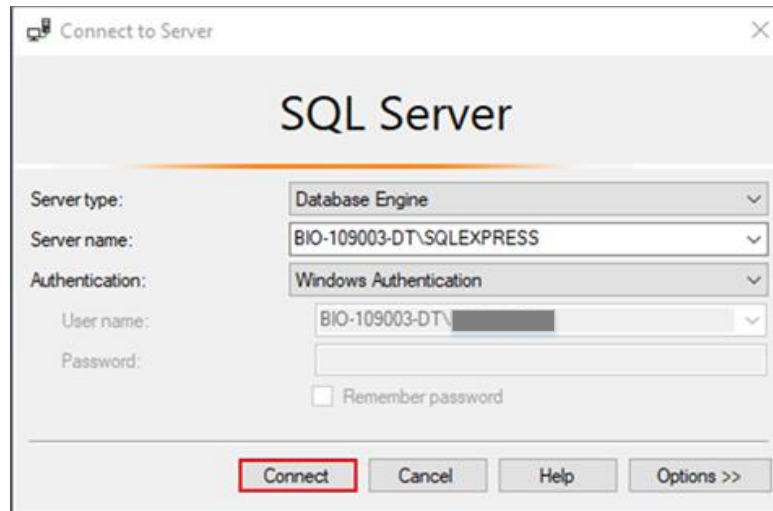



Figure 27. SQL Server Management Studio Pop Up

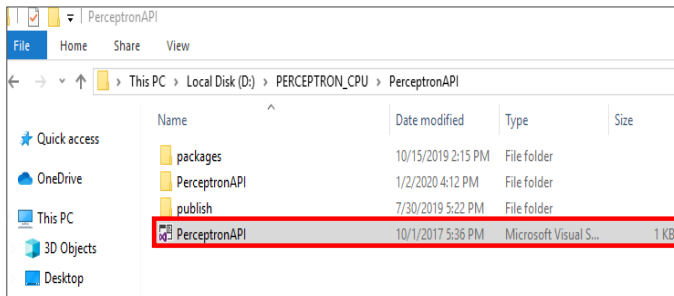
1. Click **‘Connect’**
2. Right click **‘Databases’** on left plane (Object Explorer)
3. Click **‘Restore Databases’**
4. Choose **‘Device’**, click , and add your *.bak* file
5. Click **‘OK’**, and then **‘OK’** again

10. Building PERCEPTRON

10.1 PERCEPTRON Application Programming Interface (API)

Open 'PerceptronAPI' from the folder **PERCEPTRON_CPU** in visual studio. In the Solution Explorer, click on 'PerceptronAPI'. Go to 'App_Start' and open 'WebApiConfig.cs'.

(a)



(b)

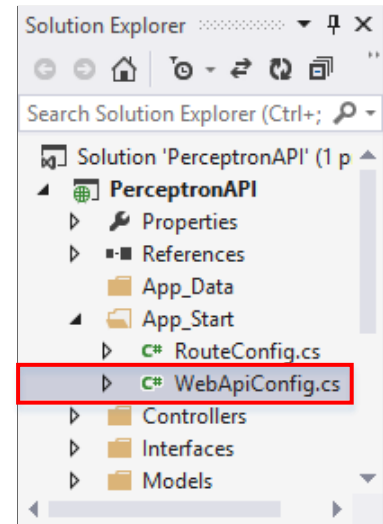


Figure 28. Building PERCEPTRON API. (a) Opening solution PerceptronAPI.sln; (b) Opening WebApiConfig.cs from Solution Explorer of Visual Studio

Next, search 'config.filters' using **CTRL+F**. Comment out the statement that includes: **config.Filters.Add(new RequireHttpsAttribute())** by typing **//** before the start of the statement.

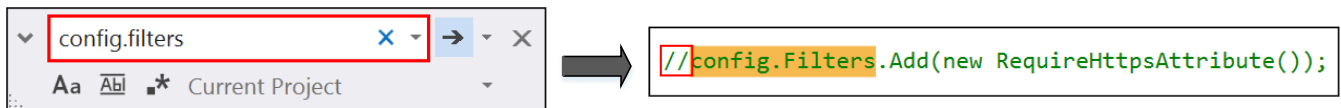


Figure 29. Searching for a statement in PERCEPTRON API solution

Open SQL Server Management Studio. Copy the given server name from SQL Server Pop Up.

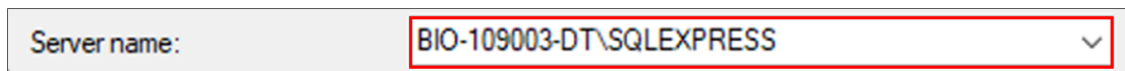


Figure 30. Copying Server Name from SQL Server Management Studio

PERCEPTRON: User Manual

In Visual Studio, open '**Web.config**' from the Solution Explorer. Press **CTRL+F** to search item. Type server name, e.g. '**CHIRAGH-II**' and press **ENTER**. Replace this server name with the name copied in Figure 30. Copying Server Name from SQL Server Management Studio, e.g. **BIO-109003-DT\SQLEXPRESS**.

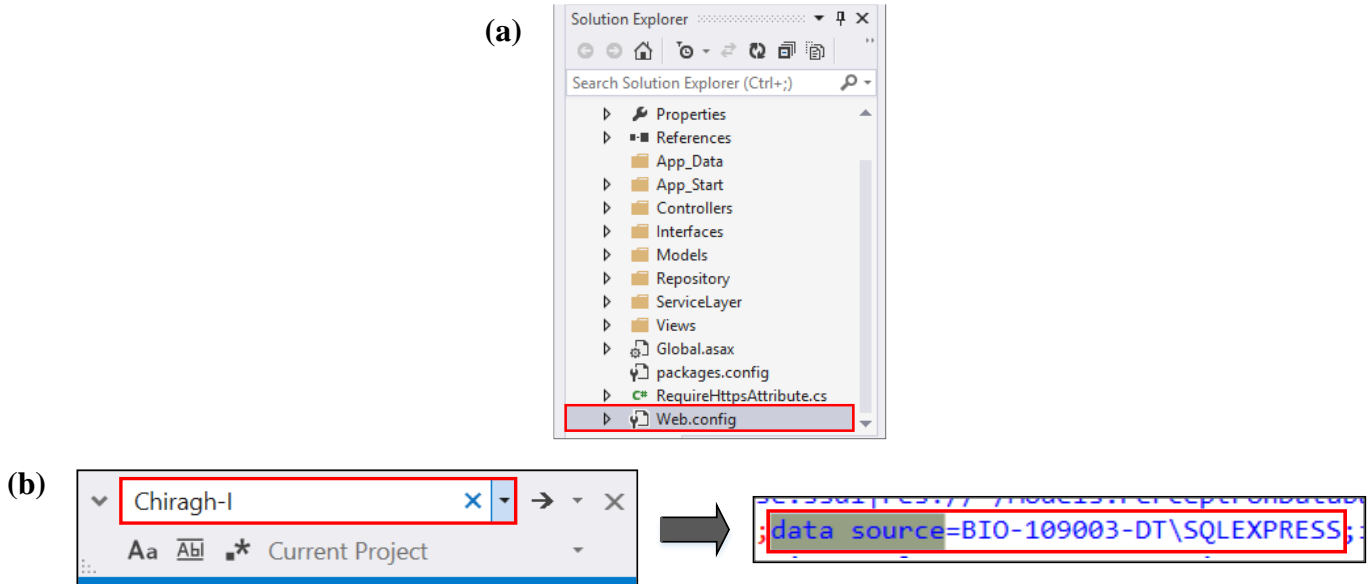


Figure 31. (a) Opening Web.config from Solution Explorer of Visual Studio (b) Changing data source (Server Name) in Web.config

Press **ENTER** again to change the name in *SqlDatabase.cs* (in Repository Folder) as well. Replace the given name for all occurrences of '**Server**' in *SqlDatabase.cs* with the server name from SQL Pop Up. In each case, add '@' right before '**Server**'.

```
using (var db = new PerceptronDatabaseEntities())
{
    var sqlConnection1 =
        new SqlConnection(
            @"Server=BIO-109003-DT\SQLEXPRESS; Database=
            PerceptronDatabase; Integrated Security=SSPI;")
}
```

Figure 32. Changing server name in *SqlDatabase.cs*

In dropdown menu of **BUILD**, select '**Build Solution.**' Next, select '**Publish PerceptronAPI**' under the same menu.

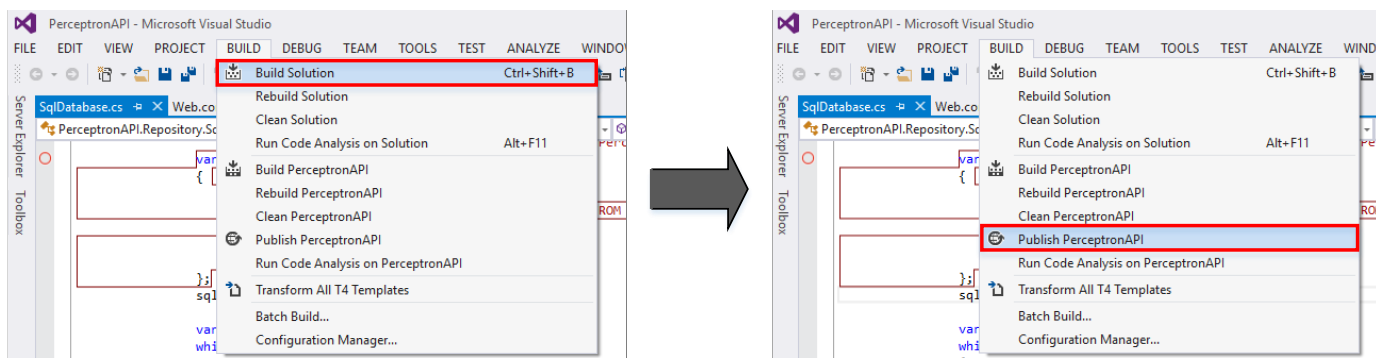


Figure 33. Building PERCEPTRON API

PERCEPTRON: User Manual

In Profile, select '**PERCEPTRON**'. In Connection, go to Target Location and ensure the following pathway: '**C:\inetpub\wwwroot\PerceptronAPI**'. In Settings, set Configuration to '**Debug**'.

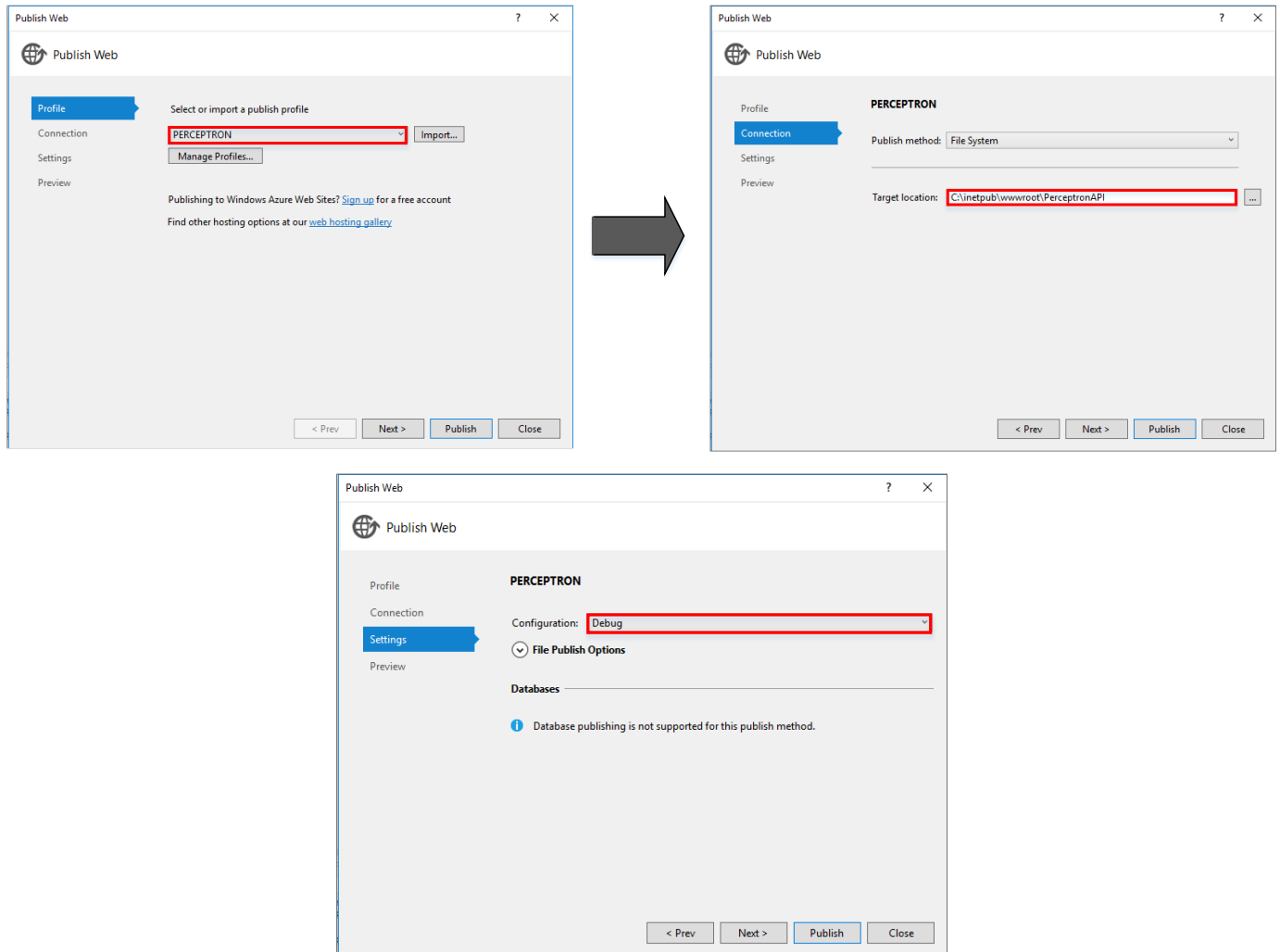


Figure 34. Steps to build PERCEPTRON API

Open the following pathway and make a folder in it by the name '**App_Data**'.

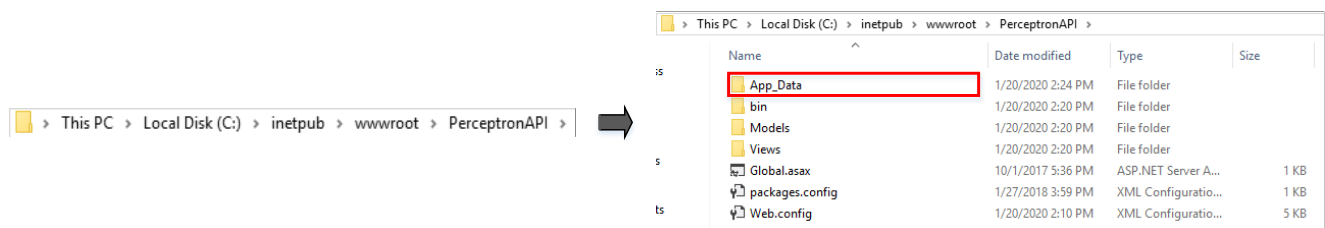


Figure 35. Publishing Perceptron API

PERCEPTRON: User Manual

From the dropdown menu of **BUILD**, select '**Build Solution.**' Run on '**Google Chrome**' and save the URL obtained.

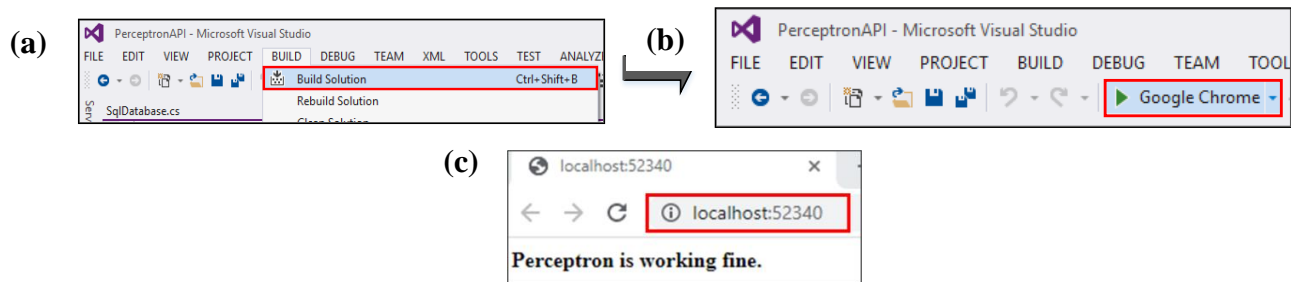


Figure 36. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine”

Message in Figure 36. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine” i.e. “PERCEPTRON API is working fine” ensures that PERCEPTRON API has been successfully built and is running fine.

10.2 PERCEPTRON Front-End

Open Visual Studio Code. From '**File**' dropdown menu, select '**Open Folder**'. Select the folder '**PerceptronFrontEnd**' and click on '**Select Folder**.'

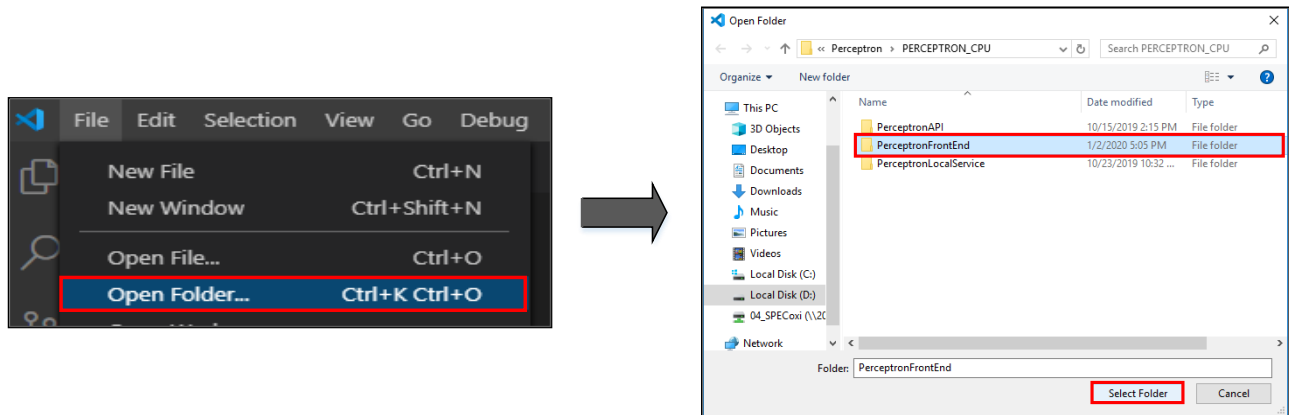
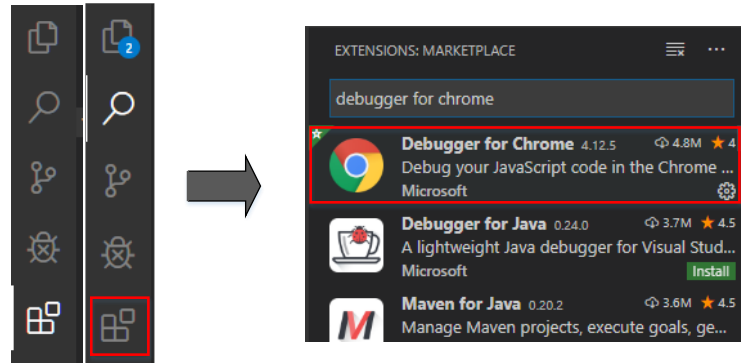


Figure 37. Opening PerceptronFrontEnd folder in Visual studio code

10.3 Connecting PERCEPTRON API to PERCEPTRON Front-End

In Visual Studio Code, click on the ‘**Extensions**’ icon (as indicated below) and type ‘**Debugger for Chrome**’ in the search bar. Select the one shown below.



Click ‘**Install**’ to install the debugger.

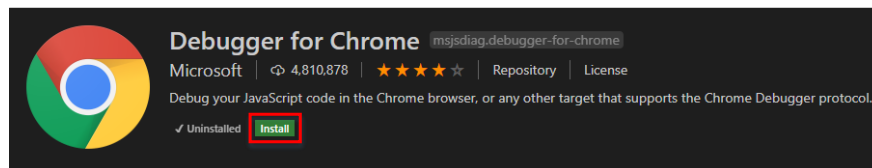


Figure 38. Installing extension in VS Code

Next, type ‘**baseAPI**’ in the search bar and select the one shown below. Replace the given URL with the one saved before in Figure 36. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine”



Type ‘**ng build**’ in Terminal. A new folder by the name of ‘**dist**’ will form in the same directory.

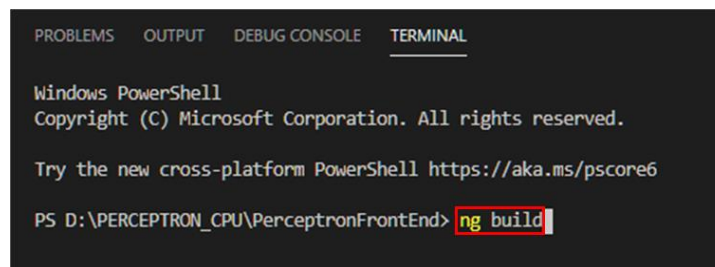


Figure 39. Building Perceptron Front-End

At bottom in Terminal Window, click on the icon  located at right side to clear Terminal and then open ‘**New Terminal**’ from **Toolbox**. Type ‘**ng serve**’ and press **ENTER**. Open the link (<http://localhost:4200/>)

PERCEPTRON: User Manual

provided in Figure 36. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine” by pressing CTRL button + left mouse button.

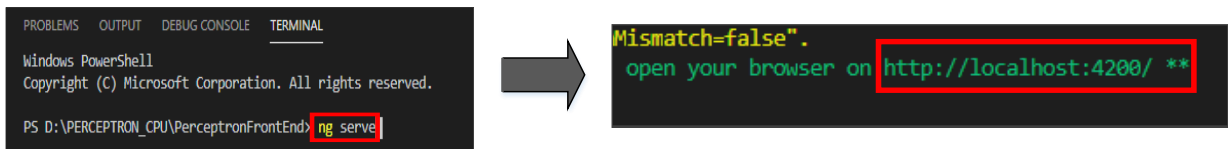
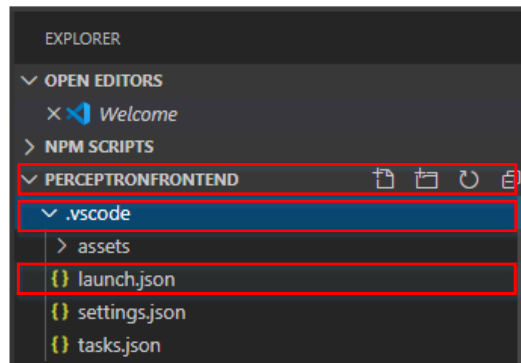
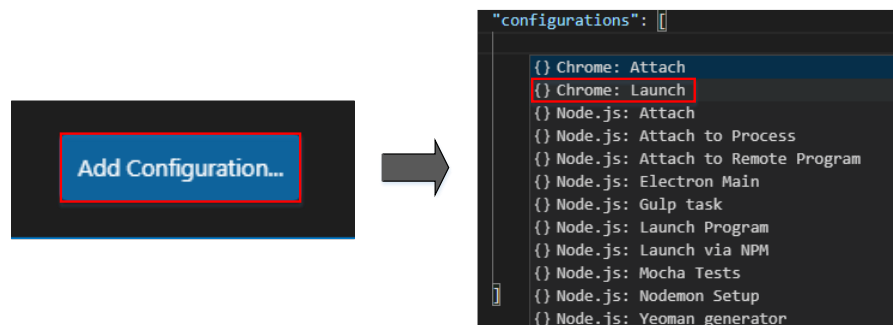


Figure 40. Hosting Perceptron Front-End on system with default URL

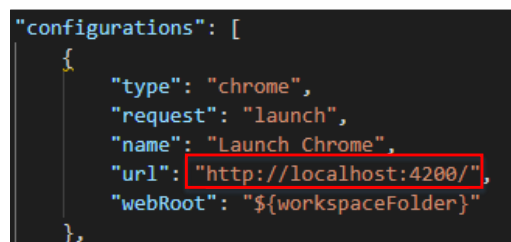
In Explorer, select ‘PERCEPTRONFRONTEND’. Click on ‘.vscode’ and open ‘launch.json’.



Click ‘Add Configuration’ icon on the bottom right corner and select ‘Chrome: Launch’ from the dropdown menu that appears



Replace the URL acquired in Figure 40. Hosting Perceptron Front-End on system with default URL.



In the Toolbox, go to ‘Debug’ and select ‘Start Debugging’. In the Pop Up that appears, click ‘Debug Anyway’ and then ‘Open launch.json’.

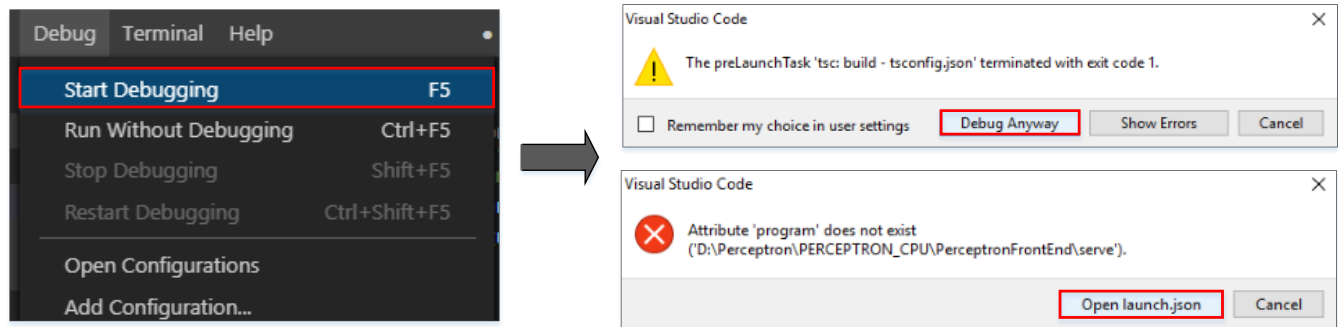


Figure 41. Steps to build PERCEPTRON Front-End

10.4 PERCEPTRON Local Service

Open SQL Server Management Studio. Copy the given server name from SQL Server Pop Up.

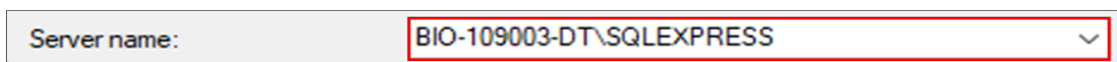


Figure 42. Copying Server Name from SQL Server Management Studio

Open 'PerceptronLocalService.sln' from the folder PERCEPTRON_CPU in visual studio. In the Solution Explorer, select 'App.config'.



Figure 43. Building Perceptron Local Service. (a) Opening PerceptronLocalService.sln; (b) Opening App.config from Solution Explorer of Visual Studio

Press **CTRL+F** to search server name. Type '**CHIRAGH-II**' and press **ENTER**. Replace this server name with the name copied in Figure 42. Copying Server Name from SQL Server Management Studio, e.g. **BIO-109003-DT\SQLEXPRESS**. Replace the given name for all occurrences of '**Data Source**' in App.config with the server name from SQL Pop Up.

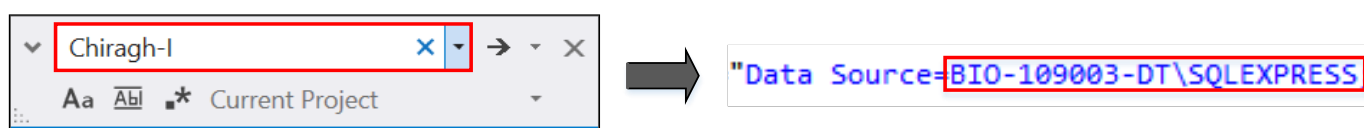
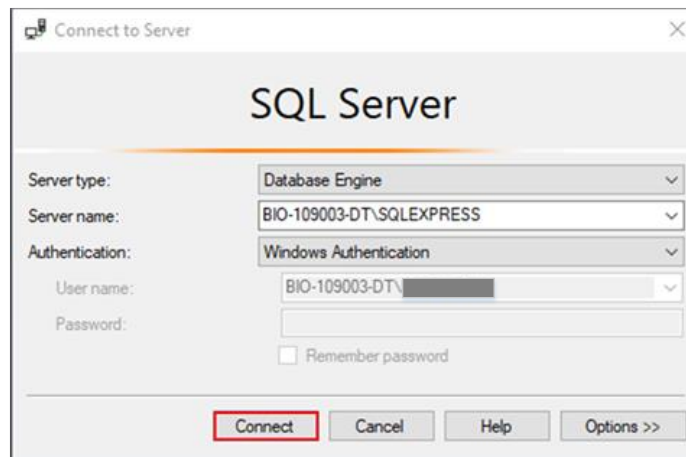


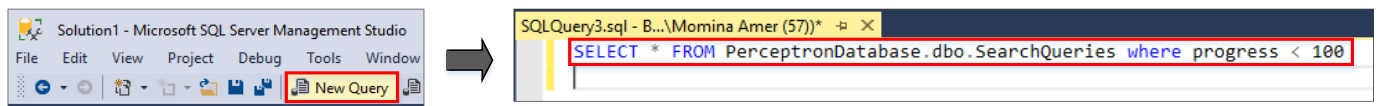
Figure 44. Changing data source value (Server Name) in App.config

Send a Protein Search Query on local host (see Section 7.2)

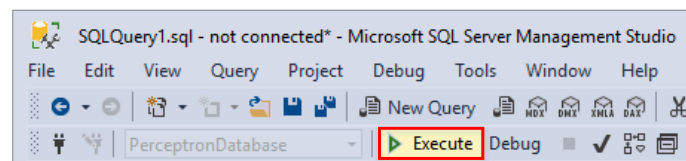
Open SQL Server Management Studio and click on '**Connect**'.



Select 'New Query' and type '**SELECT * FROM PerceptronDatabase.dbo.SearchQueries where progress < 100**' in the window that appears.



In SQL, click on 'Execute'.

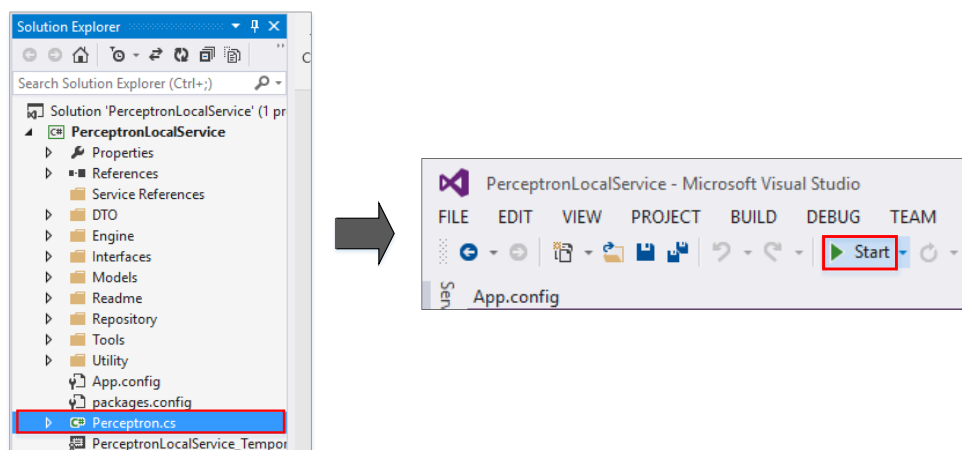


The query will show up in the Results Window.

Results					
	QueryId	UserId	Progress	CreationTime	Email
1	1ec2fba3-d044-43fc-ba87-15edc62fdd6d		0	01/21/2020 18:55:20	NULL

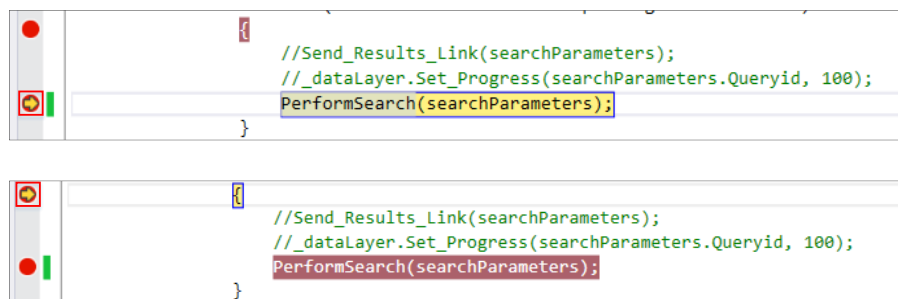
Figure 45. Steps to execute Protein Search Query

In Visual Studio, open 'PerceptronLocalService.sln'. Go to Solution Explorer and select 'Perceptron.cs'. Click 'Start' to run it.



PERCEPTRON: User Manual

Click **‘Continue’** twice to step over the breakpoints.



On the locally hosted PERCEPTRON, click on **‘Search Results & History’** to see the results of your query.

Sr. #	File Name	Protein ID	Protein Score	Molecular Weight	Truncation	Spectral Matches	No. of Modifications	Total Search Time
1	VApp_Data\HEL A_pk13_sw1_6 6sc_mono.txt	Q4VC39	0.2316	11397.7953	No	1	1	

Figure 46. Testing Perceptron Local Service

11. References

1. Pedrioli PGA, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B, Pratt B, Nilsson E, Angeletti RH, Apweiler R. A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol* 2004;22(11):1459–1466.
2. Turewicz M, Deutsch EW. Spectra, chromatograms, Metadata: mzML-the standard data format for mass spectrometer output. In: *Data mining in proteomics*. Springer; 2011. p 179–203.
3. Martens L, Chambers M, Sturm M, Kessner D, Levander F, Shofstahl J, Tang WH, Römpp A, Neumann S, Pizarro AD. mzML—a community standard for mass spectrometry data. *Mol Cell Proteomics* 2011;10(1):R110. 000133.
4. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;20(18):3551–3567.
5. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* 2012;30(10):918.
6. Liu X, Inbar Y, Dorrestein PC, Wynne C, Edwards N, Souda P, Whitelegge JP, Bafna V, Pevzner PA. Deconvolution and database search of complex tandem mass spectra of intact proteins a combinatorial approach. *Mol Cell Proteomics* 2010;9(12):2772–2782.